

Discrimination analytique des génomes bactériens Olivier Poirion

► To cite this version:

Olivier Poirion. Discrimination analytique des génomes bactériens. Autre. Ecole Centrale de Lyon, 2014. Français. NNT: 2014ECDL0033 . tel-01174161

HAL Id: tel-01174161 https://theses.hal.science/tel-01174161

Submitted on 8 Jul 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE L'UNIVERSITÉ DE LYON délivrée par l'École Centrale de Lyon

pour l'obtention du grade de

DOCTEUR

Spécialité "Ingénierie pour le Vivant"

École Doctorale Électronique, Électrotechnique, Automatique de Lyon

Soutenue publiquement le 28 novembre 2014 par

M. OLIVIER POIRION

Discrimination analytique des génomes bactériens

DEVANT LE JURY COMPOSÉ DE :

Dr David LANE (LMGE, Toulouse)	Président
Dr Éric BAPTESTE (Évolution Paris Seine)	Rapporteur
Dr Christophe DESSIMOZ (University College, London)	Rapporteur
Dr Philippe BESSIÈRES (MIG-INRA, Jouy en Josas)	Examinateur
Dr Xavier NESME (Écologie Microbienne, Lyon)	Examinateur
Dr Hervé PHILIPPE (CTMB, Moulis)	Examinateur
Dr Laurent KRÄHENBÜHL (Ampère, Écully)	Directeur de thèse
Dr Bénédicte LAFAY (Ampère, Écully)	Direction scientifique

Laboratoire Ampère CNRS UMR5005-ECL-INSA-UCBL

Je tiens à remercier mes encadrants de thèse : Bénédicte Lafay et Laurent Krähenbühl, les membres de mon jury, ma famille et, enfin, mes amis.

Résumé

Le génome bactérien est classiquement pensé comme constitué de "chromosomes", éléments génomiques essentiels pour l'organisme, stables et à évolution lente, et de "plasmides", éléments génomiques accessoires, mobile et à évolution rapide. La distinction entre plasmides et chromosomes a récemment été mise en défaut par la découverte dans certaines lignées bactériennes d'éléments génomiques intermédiaires, possédant à la fois des caractéristiques de chromosomes et de plasmides. Désignés par le terme de "chromosomes secondaires", "mégaplasmides" ou "chromids", Ces éléments sont dispersés dans l'ensemble des génomes bactériens et sont couramment décrits comme des plasmides adaptés et modifiés. Cependant, leur véritable nature et les mécanismes permettant leurs intégration dans le génome stable reste à caractériser. En utilisant les protéines liées aux Systèmes de Transmission de l'Information Génétique comme variables descriptives des éléments génomiques bactériens (ou réplicons), une étude globale de génomique comparative a été conduite sur l'ensemble des génomes bactériens disponibles. À travers l'analyse de l'information contenue dans ce jeu de données par différentes approches analytiques, il apparaît que les STIG sont des marqueurs pertinents de l'état d'intégration des réplicons dans le génome stable, ainsi que de leur origine évolutive. Les Réplicons Extra-Chromosomiques Essentiels sont de plus caractérisés comme marqueurs témoignant de la diversité et de l'évolution des mécanismes génétiques permettant l'intégration de réplicons dans le génome stable, attestant ainsi de la continuité du matériel génomique.

Summary

The genome of bacteria is classically separated into essential, stable and slow evolving replicons (chromosomes) and accessory, mobile and rapidly evolving replicons (plasmids). This paradigm is being questioned since the discovery of extra-chromosomal essential replicons (ECERs), be they called megaplasmids, secondary chromosomes or chromids, which possess both chromosomal and plasmidic features. These ECERs are found in diverse lineages across the bacterial phylogeny and are generally believed to be modified plasmids. However, their true nature and the mechanisms permitting their integration within the sable genome are yet to be formally determined. The relationships between replicons, with reference to their genetic information inheritance systems (GIIS), were explored under the assumption that the inheritance of ECERs is integrated to the cell cycle and highly constrained in contrast to that of standard plasmids. A global comparative genomics analysis including all available of complete bacterial genome sequences, was performed using GIIS functional homologues as parameters and applying several analytical procedures. GIIS proved appropriate in characterizing the level of integration within the stable genome, as well as the origins, of the replicons. The study of ECERs thus provides clues to the genetic mechanisms and evolutionary processes involved in the replicon stabilization into the essential genome and the continuity of the genomic material.

Table des matières

Remerciements	i
Résumé	iii
Summary	v
Table des matières	vii
Liste des Figures	xiii
Liste des Tables	xv
Abréviations	xvii
Lexique	xix
Notations	xxi

Introduction

1	Div	ersité	de l'architecture génomique chez les bactéries	5
	1.1	Archit	cecture génomique des trois domaines du vivant	5
	1.2	Archit	cecture génomique bactérienne : les réplicons	6
		1.2.1	Réplicons essentiels et accessoires	7
		1.2.2	Distribution des réplicons	8
		1.2.3	Taille des réplicons	8
		1.2.4	Ploïdie/multiplicité des réplicons	8
		1.2.5	Topologie des réplicons	9
		1.2.6	Spectre d'hôte des plasmides	9
		1.2.7	Rôle des plasmides	9
	1.3	Archit	ecture des réplicons : mécanismes structuraux	10
		1.3.1	Protéines s'associant au nucléoïde	10
		1.3.2	Motifs structurels du nucleoïde	13
			1.3.2.1 NAP et motifs structurels chez les plasmides	15
	1.4	Réplic	ation des réplicons	16

1

		1.4.1 Origine et initiation de la réplication
		1.4.1.1 Structure des origines de réplication
		1.4.1.2 Déroulement de l'initiation
		1.4.1.3 Régulation de l'initiation
		1.4.2 Élongation
		1.4.3 Terminaison de la réplication et résolution de dimère de réplicon . 24
		1.4.3.1 Mécanismes de terminaison de la réplication
		1.4.3.2 Résolution de dimère au niveau du site de terminaison 25
		1.4.3.3 Décaténation des régions terminales par les topoisomérases 26
	1.5	Partition active des réplicons
	1.6	Maintenance des réplicons et intégration dans le cycle cellulaire
	1.7	Mécanismes moléculaires de recombinaison, d'intégra-tion et de transfert
		génétique
		1.7.1 Résolvases impliquées dans la résolution de dimère de réplicon 32
		1.7.2 Les autres recombinases
		1.7.3 Les éléments transposables 34
		1.7.4 Conjugaison des éléments génomiques
	1.8	Fluidité du matériel génétique
		Ŭ ·
2	Les	réplicons extra-chromosomiques essentiels 39
	2.1	Découverte et premières caractérisations de génomes multipartites bactériens 3
	2.2	Diversité actuelle des génomes multipartites bactériens
	2.3	Essentialité
	2.4	Régulation et intégration des RECE dans le cycle cellulaire 45
	2.5	Origine évolutive
	2.6	$R\hat{o}le$
	2.7	Critères d'identification des RECE 55
		2.7.1 Modèle du " <i>chromid</i> " \ldots 55
		2.7.2 Autres modèles $\ldots \ldots 54$
	2.8	Données génomiques de notre étude
	2.9	Nature, origine et fonctionnement des RECE : une problématique ouverte 57
3	Str	tégie d'étude 59
Ŭ	3.1	Les protéines des STIG : variables explicatives des réplicons
	3.2	Fondement de la génomique comparative
	3.3	Concept du génome-cœur
	3.4	Fouille de données en génomique
		3.4.1 Méthodes analytiques
		3.4.2 Notations
		3.4.3 Distances 64
		3.4.4 Méthodes de clustering
		3.4.5 Méthodes de classification
		3.4.6 Projection
		3.4.7 Évaluation 66
		3.4.7.1 Performance des classifieurs
		3.4.7.2 Performance des algorithmes de clustering

			Les critères de validation externe [Gan et al., 2007; Han	
			et al., 2012])(
			Les criteres de validation interne [Gan et al., 2007; Rendon	20
			$et al., 2011] \dots \dots$)ð 20
		940	Choix des parametres)9 20
		3.4.8	Fleau de la dimensionalite)9 70
		3.4.9	Graphes	′0 - 1
	~ ~	3.4.10	Sources de blais	1
	3.5	Pipeli	ne analytique de notre étude 7	1
4	Cor	nstruct	tion de clusters de protéines homologues des STIG 7	'3
	4.1	Récup	pération des données brutes : les proté ines des STIG des génomes $\ 7$	73
		4.1.1	Principales sources publiques de protéines annotées	74
		4.1.2	Construction de la base de données requête	75
		4.1.3	Récupération des séquences protéiques à partir de RefSeq 7	76
		4.1.4	Sources de biais possibles	76
		4.1.5	Recherche d'homologues	77
	4.2	Réalis	ation de clusters d'homologues protéiques et fonctionnels 8	30
		4.2.1	Clustering des protéines	30
			4.2.1.1 Définition	30
			4.2.1.2 Principe	31
			4.2.1.3 TRIBE-MCL	31
		4.2.2	Identification des domaines fonctionnels des protéines 8	32
		4.2.3	Critères d'évaluation des clusters	32
		4.2.4	Génération de clusters aléatoires	35
			4.2.4.1 Variables	36
			4.2.4.2 Loi de probabilité des clusters aléatoires 8	37
			4.2.4.3 Vérification par test d'hypothèses	37
		4.2.5	Protocole analytique	37
			4.2.5.1 Choix d'un qr de travail	37
		4.2.6	"Nettoyage" des clusters protéiques	<i>)</i> 0
			4.2.6.1 Procédure de nettoyage	<i>)</i> 1
			4.2.6.2 Estimation de x_{seuil})1
			4.2.6.3 Résultats et discussion	<i>)</i> 2
			4.2.6.4 Alternatives à la procédure de cleaning) 4
5	Sép	aratio	n des réplicons 9)7
	5.1^{-1}	Jeux o	de données obtenues et notations	<i>)</i> 7
		5.1.1	Classification taxonomique	<i>)</i> 7
		5.1.2	Description des réplicons	<i>)</i> 7
		5.1.3	Groupes structuraux et taxonomiques)8
		5.1.4	Dimension des données)9
	5.2	Métho	odes d'évaluation de la séparation des réplicons)0
		5.2.1	Critères de validation externe utilisés)0
			5.2.1.1 V-measure)]
		5.2.2	Critères de validation interne)2
			5.2.2.1 Coefficient silhouette)2

			5.2.2.2 Critère de stabilité	. 102
		5.2.3	Sélection de modèle	. 103
			5.2.3.1 Critères	. 103
			5.2.3.2 Sélection de modèle pour les méthodes de projection	. 104
	5.3	Visual	lisation des données	. 104
		5.3.1	Réduction de dimension des données par projection	. 105
			5.3.1.1 Méthodes utilisées	. 105
			5.3.1.2 Logiciels utilisés	. 106
			5.3.1.3 Paramètres des analyses	. 106
			5.3.1.4 Projection de V^R et de \bar{V}^R_{array}	. 107
		5.3.2	Graphes	. 110
			5.3.2.1 Graphes : résultats	. 111
			5.3.2.2 Logiciels utilisés	. 116
		5.3.3	Visualisation : discussion	. 116
	5.4	Classi	fication non-supervisée	. 116
		5.4.1	Algorithmes de clustering	. 116
			5.4.1.1 Logiciels utilisés	. 119
		5.4.2	Résultats du clustering	. 120
			0	
6	Ana	alyse fo	onctionnelle des réplicons	123
	6.1	Jeux o	de données et notation	. 123
		6.1.1	Notations	. 123
		6.1.2	Dimension des données	. 124
	6.2	Discri	mination fonctionnelle des réplicons	. 125
	6.3	Discri	mination fonctionnelle des génomes	. 130
	6.4	Analy	ses par régression et test d'hypothèses	. 133
		6.4.1	Régression logistique : principe	. 133
		6.4.2	Jeux de données	. 134
		6.4.3	Résultats et discussion	. 135
		6.4.4	Conclusion	. 139
7	Clas	ssificat	tion supervisée	141
	7.1	Jeux o	de données	. 141
	7.2	Algori	ithmes de classification	. 142
	7.3	Procé	dures	. 144
		7.3.1	Classification des plasmides	. 144
		7.3.2	Classification des chromosomes	. 145
		7.3.3	Classification des génomes	. 145
	7.4	Sélecti	ion des algorithmes et des paramètres	. 145
	7.5	Logici	els utilisés	. 147
	7.6	Résult	tats et discussion	. 149
	7.7	Les no	Duveaux RECE	. 154
8	Ana	alyses	de synténie des génomes mono- et multipartites	157
	8.1	Analy	ses de synteme des genomes multipartites	. 157
		8.1.1	Protocole analytique et outils utilisés	. 158
		8.1.2	Demarche de l'étude	. 158

		8.1.3 Indice synténique	158
	8.2	Génomes des Alphaprotéobactéries	159
		8.2.1 Analyse des Rhodobactérales	159
		8.2.2 Analyse des Caulobactérales et des Sphingomonadales	164
		8.2.2.1 Caulobactérales	164
		8.2.2.2 Sphingomonadales	166
		8.2.3 Analyse des Rhizobiales	169
		8.2.4 Analyse des Rhodospirillales	174
	8.3	Génomes des Bêtaprotéobactéries	175
	8.4	Génomes des Gammaprotéobactéries	180
	8.5	Génomes des Cyanobactéries	184
		8.5.1 Analyse des Nostocales	184
		8.5.2 Analyse des Chroococcales	185
	8.6	Génomes des Bacteroidetes	186
	8.7	Discussion	188
9	Dis	cussion générale 1	191
	9.1	Méthodologies et principaux résultats	191
		9.1.1 Implications de la discrimination des RECE par les STIG	192
		9.1.2 Études complémentaires	193
	9.2	Remise en cause des hypothèses sur la nature des chromosomes secondaires	194
	9.3	Origine des biais de distribution des gènes des STIG	195
	9.4	Proposition d'un modèle moléculaire d'origine des néo-chromosomes	196
	9.5	Continuité du matériel génomique	197
Bi	bliog	graphie 1	197
A	nnex	es 2	227
A	Div	ersité des réplicons secondaires accessoires	229
В	Gro	oupes d'orthologie KEGG sélectionnés	237
С	Fan	nilles protéiques ACLAME sélectionnées	241
D	Rés	ultats du clustering de V^R par INFOMAP	245

Table des figures

1.1	Diversité des réplicons des trois domaines du vivant 6
1.2	Représentation schématique de la réplication du chromosome bactérien 16
1.3	Structure de l'origine de réplication de <i>E. coli</i>
1.4	Diversité structurale des origines de réplication chez les bactéries 20
2.1	Distribution des génomes multipartites dans le domaine Bacteria 41
2.2	Distributions des tailles et nombres de gènes des réplicons
2.3	Répartition par lignée des génomes bactériens disponibles
3.1	Pipeline analytique
4.1	Procédure de récupération des données brutes
4.2	Choix des termes KEGG BRITE
4.3	Procédure de clustering des protéines
4.4	Distribution des types et nombres d'occurence des profiles dans P_{homo} 83
4.5	Influence de la granularité sur le clustering
4.6	Distribution de la taille des clusters de protéines pour $gr = 4$
4.7	Pourcentage de l'annotation la plus fréquente par cluster 90
4.8	Annotation des clusters de protéines
4.9	Différents niveaux d'homologie chez les tyrosine-recombinases 95
5.1	Projection de l'ensemble des réplicons
5.2	Projection des réplicons normés par genre bactérien
5.3	Identification des lignées bactériennes sur la projection des réplicons 110
5.4	Visualisation des réplicons par graphe bipartite
5.5	Visualisation par graphe des réplicons par lignée bactérienne
5.6	Pourcentages de connexion aux clusters de protéines par type de réplicon 115
6.1	Variance expliquée par les quatre composantes principales
6.2	Projections des réplicons selon les quatres composantes principales d'une
	ACP
6.3	Visualisation des lignées bactériennes sur la projection des réplicons selon les deux composantes principales d'une ACP
6.4	Projection fonctionnelle des génomes selon les deux composantes princi- pales d'une ACP
8.1	Synténie de Paracoccus vs. R. sphaeroides
8.2	Synténie de Rhodobacter vs. D. shibae
8.3	Synténie de Ruegeria vs. D. shibae

8.4	Synténie d'Asticca caulis $vs.$ autre Caulobacteraceae et Sphingomonaceae . 165
8.5	Synténie de Sphingobium vs. Caulobacteraceae et autres Sphingomona-
	daceae
8.6	Hypothèse d'évolution du plasmide ITR chez les Rhizobiales
8.7	Synténie de Brucella vs. autres Rhizobiales
8.8	Synténie de Azospirillum vs. Rhodospirillum
8.9	Synténie de Burkholderia vs. autres Burkholderiales
8.10	Synténie de Ralstonia vs. autres Burkholderiales
8.11	Synténie de Aliivibrio vs. gammaprotéobactéries proches
8.12	Synténie entre espèces multi-/monopartite d'Anabaena
8.13	Synténie entre espèces multi-/monopartite de Cyanothece
8.14	Synténie de <i>Prevotella vs.</i> autres Bacteroidetes
9.1	Synthèse des résultats

Liste des tableaux

1.1	Comparaison des structures génomiques des trois domaines du vivant	6
1.2	Principales protéines s'associant au nucléoïde (NAP)	11
1.3	Propriétés des NAP majoritaires des bactéries	12
1.4	Motifs structurels du nucleoïde	13
1.5	Caractéristiques des principaux motifs trouvés chez les bactéries	15
1.6	Principales protéines impliquées dans l'initiation de la réplication.	17
1.7	Principales protéines impliquées dans l'élongation.	23
1.8	Principales protéines impliquées dans la terminaison de la réplication.	25
1.9	Principales protéines impliquées dans la ségrégation des réplicons.	27
1.10	Principales protéines impliquées dans la maintenance et le cycle cellulaire	28
1.11	Résolvases impliquées dans la résolution de dimère	32
1.12	Les recombinases distinctes des résolvases.	33
1.13	Principaux types de transposons	34
2.1	Caractéristiques essentielles des RECE	43
2.2	Origine évolutive des RECE d'après la littérature	48
4.1	Annotations utilisées pour la selection des Familles ACLAME	76
4.2	Principaux algorithmes d'alignement de séquences	78
4.3	Principales annotations multiples identifiées parmi les clusters	89
5.1	Dimension des données utilisées	99
5.2	Ensembles des classes de référence Kl_{ref} selon la séparation étudiée 1	01
5.4	Méthodes de projection utilisées	05
5.5	Évaluation des procédures de visualisation de l'ensemble des réplicons \dots 1	07
5.6	Algorithmes de clustering utilisés	17
5.7	Évaluation des procédures de clustering de V^R et \bar{V}^R_{qenre}	19
5.8	Classification non-supervisée des RECE	22
6.1	Dimension des données fonctionnelles	25
6.2	Évaluation des procédures de clustering fonctionnel des réplicons 1	26
6.3	Classification non-supervisée des RECE	30
6.4	Évaluation du clustering fonctionnel des génomes	31
6.5	Propriétés des classes utilisées pour l'analyse de régression logistique 1	35
6.6	Résultats des régressions logistiques entre les différentes classes d'éléments	
	génomiques	36
6.7	Résultats des régressions logistiques entre les différentes classes de génomes1	38
7.1	Taille des ensembles formant les <i>training sets</i>	42

7.2	Algorithmes de classification utilisés
7.3	Comparaison des classifieurs pour la classification supervisée des réplicons 147
7.4	Plasmides classés comme RECE
7.5	Probabilités des RECE d'appartenir à la classe "RECE"
7.6	Réplicons extra-chromosomiques classés comme chromosome
7.7	Importance des attributs fonctionnels des observations de $E_{training}$ dans
	la classification RECE/plasmide
7.8	OOB_{score} obtenus avec ERT et les clusters proté iques en tant qu'attributs 152
7.9	Probabilités des génomes multipartites d'appartenir à la classe "multipar-
	tite"
8.1	Valeurs de l'indice synténique pour <i>Paracoccus</i>
8.2	Valeurs de l'indice synténique pour <i>Rhodobacter</i>
8.3	Valeurs de l'indice synténique pour <i>Ruegeria</i>
8.4	Valeurs de l'indice synténique pour Asticcacaulis
8.5	Valeurs de l'indice synténique pour <i>Sphingobium</i>
8.6	Valeurs de l'indice synténique pour Brucella
8.7	Valeurs de l'indice synténique pour Azospirillum
8.8	Valeurs de l'indice synténique pour <i>Burkholderia</i>
8.9	Valeurs de l'indice synténique pour Ralstonia
8.10	Valeurs de l'indice synténique pour Aliivibrio
8.11	Valeurs de l'indice synténique pour Anabaena
8.12	Valeurs de l'indice synténique pour <i>Cyanothece</i>
8.13	Valeurs de l'indice synténique pour <i>Prevotella</i>
A.1	Diversité des réplicons secondaires essentiels
B.1	Groupes d'orthologie de KEGG sélectionnés
C.1	Familles ACLAME liées à la réplication
C.2	Familles ACLAME liées à la ségrégation
C.3	Familles ACLAME liées à la résolution de dimères
C.4	Familles ACLAME liés à la maintenance

Abréviations

ARNr	ARN ribosomique
ARNt	ARN de transfert
chr	chromosome
ICE	Integrative Conjugative Element, $(\S1.13)$
IME	Integrative Mobile Element, $(\S1.13)$
IPP	Interaction Protéines-Protéines
ITR	(plasmide) Intragenomic Tanslocation Recipient ($\S8.2.3$)
NAP	Nucleoid Associated Protein
MGE	Mobile Genetic Element, $(\S1.13)$
ori	Origine de réplication $(\S1.2)$
oriT	Origine de transfert $(\S1.7.4)$
pb	Paire de bases
OR	$\mathbf{O}\mathrm{dd}\ \mathbf{R}\mathrm{atio},$ rapport de chance, défini selon un modèle de régression
	logistique $(\S6.4.1)$
PSK	P ost- S egregational K illing (§1.6)
\mathbf{RC}	Rolling Circle (replication)
RECE	\mathbf{R} éplicon Extra-Chromosomique Essentiel
\mathbf{SD}	Strand Displacement (replication)
STIG	\mathbf{S} ystèmes de Transmission de l'Information Génétique
THG	Transfert Horizontal de Gène, $(\S1.8)$

Lexique

Observation	Désigne un élément extrait d'un jeu de données		
Attribut	Caractéristique décrivant une observation		
Clustering	Désigne un processus de partition d'un ensemble d'observations		
Génome-coeur	Fraction du génome partagée par tous les membres d'une espèce (§3.3)		
Fitness	Désigne ce qui a rapport avec les capacités de survie et de développement d'un organisme biologique à court, moyen et long terme, dans ses habitats naturels ($\S2.3$)		
Data mining	Fouille de données. Ensembles de méthodes analytiques (§3.5)		
Input	Donnée(s) d'entrée d'un processus analytique ou algorithmique		
Outlier	Donnée d'un jeu de données, distantes de la majorité des données		
Output	Donnée(s) de sortie d'un processus analytique ou algorithmique		
Overfitting	Sur-apprentissage $(\S3.5)$.		
Machine learning	Ensemble de méthodes analytiques fondées sur l'apprentissage (§3.5)		
Training set	Jeu de données d'apprentissage avec le quel un algorithme peut être entraı̂né avant d'être testé sur un jeu de données $(\S 3.5)$		

Notations

R	Désigne l'ensemble des réplicons considérés dans l'étude
G	Désigne l'ensemble des génomes considérés dans l'étude
Kl	Désigne un clustering d'un ensemble de réplicons ou de génomes
Cl	Désigne un clustering d'un ensemble de protéines
V^R	Désigne un ensemble de vecteurs où chaque vecteur correspond à un réplicon de R ayant comme attributs les protéines codées par celui-ci et appartenant à des clusters d'homologues (§5.1)
$ar{V}^R_{genre}$	Désigne un ensemble de vecteurs où chaque vecteur correspond à un ensemble de réplicons de R du même type et du même genre taxonomique et où ce vecteur correspond au barycentre des vecteurs de réplicons de ce groupe (§5.1)
V_f^R	Similaire à V^R mais les attributs des vecteurs sont les annotations des clusters d'homologues (§6.1)
V_f^G	Similaire à V_f^R mais les observations considérées sont les génomes (§5.1)
$ar{V}^R_{f,genre}$	Similaire à \bar{V}^R_{genre} mais les attributs considérés sont les annotations des clusters d'homologues (§6.1)
$\bar{V}^G_{f,genre}$	Similaire à $\bar{V}^R_{f,genre}$ mais les observations considérées sont les génomes (§6.1)
Cl^{KEGG}	Ensemble des familles d'orthologues proté iques de KEGG sélectionnées (§4.1.2)
Cl^{ACLAME}	Ensemble des familles de protéines ACLAME sélectionnées ($\S4.1.2$)
P^{ref}	Ensemble des protéines de référence d'ACLAME et de KEGG
gr	Granularité de l'algorithme TRIBE-MCL

Introduction

Différents critères peuvent nous servir à séparer de façon naturelle ce qui est vivant de ce qui ne l'est pas :

- La structure : les organismes vivants sont organisés en une ou plusieurs cellules.
- L'information : Tout organisme vivant présente, au sein de ses cellules, de l'information qui est codée au niveau de l'ADN.
- La continuité : Au cours de sa vie, un organisme vivant doit mettre en place des mécanismes assurant la transmission conforme de son matériel génétique à sa descendance permettant ainsi sa viabilité.

Les architectures des génomes bactériens sont de même soumises à ces principes et il a été caractérisé, au sein de divers génomes bactériens, de nombreux mécanismes génétiques permettant la transmission de l'information génomique (Chapitre 1). Sur cette base, le génome est organisé en **réplicon(s)**, ou unité(s) de réplication, (§1.2) qui, chez les bactéries, correspondent à des chromosomes ou des plasmides selon leur caractère essentiel ou accessoire pour l'organisme hôte (§1.2.1). En plus de leur rôle de support de l'information génétique, les réplicons bactériens sont des éléments fluides où prennent place différents phénomènes de transfert intra- ou inter-génomique, participant à l'adaptation et l'évolution du matériel génétique et des organismes biologiques.

L'ensemble des connaissances des Systèmes de Transmission de l'Information Génétique (STIG) provient de l'étude d'un petit nombre d'espèces bactériennes et semblent montrer l'existence de caractéristiques spécifiques des STIG des chromosomes, nécessitant une stabilisation dans le cycle cellulaire, et des STIG des plasmides, pouvant permettre un comportement pseudo-autonome de ces derniers $(\S1.8)$. La conception traditionnelle du génome bactérien est qu'il est constitué d'un unique chromosome présentant des mécanismes d'intégration dans le cycle cellulaire découplés de ceux des autres réplicons additionnels présents dans la cellule. La découverte assez récente de Réplicons Extra-Chromosomiques Essentiels (RECE), réplicons structurellement différents du chromosome bien que possédant des gènes essentiels (Chapitre 2), pose alors la question des mécanismes génétiques d'intégration des RECE dans le cycle cellulaire en parallèle des chromosomes ($\S2.4$). Au delà de l'aspect mécanistique, le rôle et l'origine des RECE sont des problématiques ouvertes dont la compréhension peut apporter des éléments de réponse quant à la complexification du génome et des organismes vivants par le passage entre structure génomique monopartite, à un chromosome, et structure multipartite, à plusieurs chromosomes ($\S2.9$).

On peut faire l'hypothèse que les RECE, en tant qu'espèce génomique propre, sont dans

un état d'intégration du cycle cellulaire distinct des chromosomes et des plasmides. Les STIG, mécanismes génétiques responsable de cette intégration, devraient alors pouvoir fournir des critères de discrimination des plasmides, chromosomes et RECE (Chapitre 3). L'approche alors suivie est de caractériser les réplicons bactériens par leurs STIG.

Plus formellement, on peut modéliser l'hypothèse de la façon suivante :

Stabilisation = f(STIG)

L'objectif est alors d'inférer le modèle f décrivant la réalité et inconnu. Une première remarque est que l'on peut naturellement supposer que le modèle réel f prend en compte des paramètres λ additionnels et inconnus :

$$Stabilisation = f(STIG, \lambda)$$

Par exemple, on peut supposer que l'état de stabilisation d'un réplicon est fortement couplé à l'écologie de l'hôte bactérien (comme ce qui semble exister chez les Rhizobiales). Cependant il est raisonnable de penser que les STIG eux-mêmes reflètent ces paramètres externes : un réplicon additionnel fortement stabilisé d'un hôte symbiote aura des STIG spécifiques. Ainsi, l'utilisation des STIG est suffisante pour modéliser f.

Le travail présenté dans cette thèse a consisté à proposer des modèles \hat{f} approximés de f, déduits d'ensembles de données \widehat{STIG} liés aux STIG et Stabilisation liés à la stabilisation des génomes. On a ainsi la relation suivante :

$$Stabilisation = \hat{f}(\widehat{STIG}) + erreur$$

où l'erreur est d'autant plus faible que \hat{f} est probablement proche de f.

Les jeux de données Stabilisation sont construits à partir d'un ensemble de génomes et réplicons et de leurs annotations (chromosome, plasmide ou RECE) par les bases de données publiques dont ils sont extraits (Chapitre 4). Les données STIG utilisées sont construites à partir de clusters de protéines homologues en terme de séquence et de domaines fonctionnels, qui sont sélectionnés initialement par leurs homologies de séquence envers des protéines annotées fonctionnellement et en lien avec les STIG (Chapitre 4). Les clusters de protéines sont ensuite employés comme attributs des réplicons bactériens pour structurer l'espace des réplicons par des analyses non-supervisées de clustering et de visualisation (Chapitre 5). Les clusters de protéines sont obtenus en utilisant un ensemble d'une centaine de fonctions liées aux STIG (Annexes B et C). Ces fonctions sont ensuite directement utilisées pour caractériser les différents réplicons et génomes (mono- et multi-partites) bactériens (Chapitre 6). Ces premières études permettent de mettre en évidence des spécificités des différents types de réplicons, ce qui permet, avec l'utilisation d'algorithmes d'apprentissage supervisé, d'identifier de nouveaux RECE potentiels parmi les réplicons bactériens (Chapitre 7). Enfin, afin de mieux comprendre les mécanismes évolutifs impliqués dans la formation des génomes multipartites, des analyses complémentaires de synténie sont réalisées entre génomes multi- et monopartites (Chapitre 8). Les résultats de ces différentes analyses permettent ainsi de qualifier les différences fonctionnelles en terme de STIG entre les réplicons bactériens, et mènent à proposer des modèles évolutifs de formation des génomes multipartites et à souligner la continuité du matériel génomique bactérien (Chapitre 9).

Chapitre 1

Diversité de l'architecture génomique chez les bactéries

Ce chapitre a pour objectif de détailler le contexte biologique et génomique de l'étude à partir de la littérature existante. Nous nous focaliserons cependant sur les mécanismes de réplication/ségrégation et maintenance du cycle cellulaire. Dans un deuxième temps, nous soulèverons la question de l'étude, nous détaillerons l'importance de celle-ci dans le contexte de la génomique évolutive et des potentielles implications sur la compréhension de l'évolution des génomes bactériens.

1.1 Architecture génomique des trois domaines du vivant

Tout organisme vivant est composé de cellule(s). L'information les définissant a pour support physique un génome, ensemble du matériel génétique, composé de molécule(s) d'acide désoxyribonucléique (ADN). L'information qu'il code, nécessaire au développement et au fonctionnement cellulaire, s'exprime non seulement dans la synthèse d'acide ribonucléique (ARN) et de protéines par l'intermédiaire du code génétique, mais aussi à travers l'organisation spatiale du génome. L'information codée permet de plus sa propre reproduction et, ainsi, assure la conservation du matériel génétique à travers le temps et les générations. Enfin, le génome permet l'évolution et l'adaptation d'une population par sa capacité à être modifié de génération en génération. Tous les organismes ne possèdent pas la même architecture génomique, celle-ci reflétant notamment des différences de morphologie et d'écologie des espèces.

Les organismes vivants sont classés en trois domaines : Archées, Eucaryotes et Bactéries. Les virus, parfois proposés comme quatrième domaine du vivant [McGeoch and Bell, 2008], en sont néanmoins exclus actuellement car n'existant pas sous une forme cellulaire. Ces trois domaines se différencient par plusieurs propriétés essentielles (Table 1.1) mais possèdent néanmoins des constituants génétiques similaires ainsi que les mêmes schémas fonctionnel et organisationnel fondamentaux.

Caractéristiques	Bactéries	Archées	Eucaryotes
Génome			
Noyau	-	-	+
Gyrase inverse	-	+	-
ARN			
Polymérase simple	+	-	-
Polymérase complexe	-	+	-
Polymérases multiples	-	-	+
ARNm polycistronique	-	-	+
Ribosomes 70S	+	+	+/-
Introns			
ARNt	-	+/-	+
ARNr	+/-	+/-	-
ARNm	+/-	+/-	-

 TABLE 1.1: Comparaison des structures génomiques des trois domaines du vivant.

 Adapté de [Perry and Staley, 1997]

1.2 Architecture génomique bactérienne : les réplicons

Dans la compréhension de l'architecture des génomes, le terme de **réplicon** désigne une unité réplicative du génome. Ce terme a été introduit en 1963 [Jacob et al., 1963] et désigne de façon générale une unité de réplication indépendante se répliquant dans son ensemble. D'un point de vue structural, un réplicon est l'élément génétique défini entre une origine de réplication active et un site de terminaison (Figure 1.1).



FIGURE 1.1: Diversité des réplicons des trois domaines du vivant.

(a) Chromosome bactérien type avec une unique origine de réplication (*ori*) et un site de terminaison à l'opposé.
(b) Réplicon de type archée avec plusieurs *ori* inspiré du génome de *Sulfolobus islandicus*.
(c) Génome eucaryote typique à multiples origines de réplication ORC. Adapté de [Hyrien et al., 2013].

Différents réplicons peuvent être présents sur une même molécule d'ADN, ce qui est le cas général chez les Eucaryotes et semble être fréquent chez Archées [Samson and Bell, 2011, cette configuration permettant alors de multiples amorçages de l'initiation de la réplication. Dans ce cas, il n'existe pas de site de terminaison précis. Chez les bactéries par contre, il est communément admis que chaque molécule d'ADN possède un unique réplicon actif. Dans l'étude du génome d'Escherichia coli, un des principaux et premiers organismes modèles en biologie, l'initiation de la réplication du génome, constitué d'un unique chromosome circulaire, s'effectue à un seul site, correspondant à la séquence d'origine de réplication ou ori [Skarstad et al., 1986]. Ce phénomène, étant observé chez d'autres espèces bactériennes ainsi que pour des plasmides (voir ci-dessous) Messer, 2002; Worning et al., 2006], le terme "réplicon" a été généralisé pour désigner, chez les bactéries, les différentes entités génomiques composées d'une unique molécule d'ADN. Quelques exceptions existent toutefois dans le cas des plasmides bactériens. Par exemple, des plasmides tels que les plasmides pGSH500 et pJD4 de Klebsiella pneumoniae et Neisseria gonorrhoeae respectivement, possèdent plusieurs ori (réplicons sensu stricto), ce qui leur permet de mieux s'intégrer dans le génome de multiples hôtes [Toukdarian, 2004].

1.2.1 Réplicons essentiels et accessoires

Le génome bactérien est constitué de réplicons essentiels à la survie des cellules, les chromosomes, et de réplicons non essentiels, les plasmides [Mackenzie et al., 2004]. Historiquement, le terme de "plasmide" a été introduit par Lederberg en 1952 [Lederberg, 1952] et désignait à l'époque toute particule extra-chromosomique capable de se reproduire dans un état autonome. Cette définition englobait en fait ce que l'on désigne aujourd'hui sous les termes de "virus", "bactériophage", "ribosome", "transposon", "séquence d'insertion" en plus de la définition actuelle de "plasmide" [Lederberg, 1952]. Comment alors définir proprement un plasmide? Par opposition à un chromosome et aux autres éléments extra-chromosomiques, un plasmide peut être caractérisé par son spectre d'hôtes, son mode de réplication plus ou moins autonome vis-à-vis du génome de l'espèce hôte, sa capacité de mobilité inter-hôte, et par des mécanismes d'incompatibilité inter-plasmidique [Helinski, 2004; Slater et al., 2008]. Les plasmides diffèrent des chromosomes en présentant des très hauts taux de flux génétiques, et peuvent porter des éléments génétiques provenant de bactéries qui n'ont pas ou peu de caractéristiques en commun. Les plasmides peuvent donc être vus comme des mosaïques de modules fonctionnels aux origines phylogéniques diverses [Fernández-López et al., 2006].

La notion d'essentialité pour un réplicon est cependant ambiguë, dans le sens où un réplicon peut être qualifié d'essentiel si et seulement si il comporte au moins un gène essentiel à la survie de l'organisme [Mackenzie et al., 2004]. Des marqueurs traditionnellement utilisés pour déterminer l'essentialité d'un réplicon sont les gènes codant les ARN ribosomiques (ARNr) [Mackenzie et al., 2004]. L'étude de génomes réduits a fourni différents ensembles de gènes considérés comme étant essentiels à la survie des organismes bactériens [Glass et al., 2006; Koonin, 2000], qui peuvent être, de même, utilisés comme marqueurs d'essentialité de l'élément génétique. Mesurer l'essentialité d'un gène est cependant une tâche ardue en pratique. Elle dépend souvent de l'écologie de l'organisme et peut varier, pour un gène donné, d'un organisme à un autre. Enfin, dans certaines conditions, la perturbation d'un gène peut gravement entraver la capacité d'un organisme à se développer et à prospérer sans forcément le tuer [Mackenzie et al., 2004; Slater et al., 2008]. Doit-on alors trancher sur le statut "essentiel" ou "accessoire" d'un gène donné selon la survie de l'organisme en cas d'altération du géne, ou peut-on plutôt parler de différents degrés d'essentialité d'un gène, et par extension d'un réplicon, selon l'organisme considéré?

1.2.2 Distribution des réplicons

Le génome d'*E. coli*, espèce modèle de référence, est composé d'un large réplicon : le chromosome et, selon les souches, de réplicons plasmidiques additionnels plus petits. Ce schéma est largement retrouvé parmi les bactéries. La structure typique des génomes bactériens est donc : un unique chromosome, coexistant éventuellement avec des plasmides. Cette constatation a été largement utilisée afin de décrire les génomes bactériens : le réplicon le plus large étant généralement désigné comme le chromosome et considéré comme essentiel, et les réplicons additionnels étant considérés comme des plasmides, donc accessoires [Casjens, 1998]. Néanmoins, de nombreuses exceptions existent au sein des génomes bactériens (ces éléments seront approfondis dans le Chapitre 2), ce qui remet en cause ce paradigme.

1.2.3 Taille des réplicons

Les réplicons bactériens présentent une très grande diversité de taille. Pour les chromosomes et réplicons extra-chromosomiques dont les séquences complètes étaient disponibles dans la base de données publiques RefSeq à la date du 23 octobre 2013, la taille des chromosomes varie de 139 kb (*Candidatus* Tremblaya princeps (*PCVAL* [López-Madrigal et al., 2011]) à 14.782 kb (*Sorangium cellulosum* [Han et al., 2013]) et de 0,74 kb (*Candidatus* Tremblaya phenacola *PAVE* [Husnik et al., 2013]) à 2.580 kb (*Cupriavidus metallidurans* CH34 (NC_007974.2)) pour les plasmides. Il est intéressant de constater que la taille du plus grand des plasmides est 20 fois plus importante que celle du plus petit des chromosomes : Taille du plus grand plasmide de taille sont notablement différentes, elles ne sont pas strictement distinctes.

1.2.4 Ploïdie/multiplicité des réplicons

Les chromosomes bactériens sont souvent perçus comme haploïdes [Casjens, 1998] par opposition aux plasmides qui sont vus comme potentiellement polyploïdes. Même si ces postulats se révèlent souvent corrects, il existe de nombreuses exceptions à cette règle. *Deinococcus radiodurans* [White et al., 1999], *Epulopiscium* sp. [Mendell et al., 2008], ou *Neisseria gonorrhoeae* [Tobiason and Seifert, 2006] présentent une multiplicité de leur chromosome. Il existe de plus une grande variation du nombre de copies possibles d'un plasmide dans une cellule bactérienne [Helinski, 2004]. Pour décrire cette diversité, deux exemples peuvent être cités : les mégaplasmides de la famille des Rhizobiaceae qui, pour certains, sont présents en une unique copie dans l'organisme [Pinto et al., 2012] et le plasmide "modèle" d'*E.coli*, ColE1, qui peut être présent en plus de 20 copies dans une cellule [Summers and Sherratt, 1984]. La régulation du nombre de copies chez les plasmides et chromosomes passe par de nombreux mécanismes génétiques que nous allons aborder dans les sections suivantes.

1.2.5 Topologie des réplicons

Même si une configuration circulaire est la norme chez les réplicons plasmidiques et chromosomiques bactériens [Casjens, 1998], il existe différents exemples de chromosomes [Chaconas and Chen, 2005] et de plasmides [Stewart et al., 2004] adoptant une configuration linéaire. Par exemple, dans le génome de *Borrelia burgdoferi*, le chromosome de 0,9 Mb et 12 des 22 plasmides présentent une configuration linéaire. Divers réplicons linéaires chromosomiques (de grande taille) et plasmidiques sont aussi trouvés chez les Actinobactéries. Différentes solutions au problème de la terminaison de la réplication des réplicons linéaires coexistent dans le domaine bactérien [Hinnebusch and Tilly, 1993], suggérant des évolutions multiples indépendantes de cette topologie. Le rôle et les avantages d'une topologie linéaire des réplicons restent cependant flous [Casjens, 1998; Chaconas and Chen, 2005]. Une apparition secondaire par linéarisation de la forme circulaire est favorisée [Volff and Altenbuchner, 2000] et peut être réversible.

1.2.6 Spectre d'hôte des plasmides

Les plasmides sont différenciés selon leur capacité à exister dans divers hôtes bactériens. On distingue alors les plasmides à spectre d'hôtes étroit ou large selon le nombre d'hôtes qui peuvent les héberger. Cette propriété est principalement en rapport avec une adaptation particulière des mécanismes de la réplication, ségrégation et maintenance plasmidiques [Jain and Srivastava, 2013; Toukdarian, 2004].

1.2.7 Rôle des plasmides

À travers les différents aspects abordés, aucune distinction majeure entre chromosomes et plasmides n'apparaît, hormis un aspect essentiel ou accessoire des réplicons. Néanmoins, les chromosomes, en tant que partie essentielle du génome bactérien, *définissent* leur hôte. Ainsi, toute spécificité génomique de ceux-ci reflétera l'adaptation de l'organisme bactérien à son écologie et à son mode d'évolution [Bentley and Parkhill, 2004]. Inversement, les éléments plasmidiques peuvent présenter un caractère "égoïste" par l'évolution de stratégies permettant leur existence propre sans forcément apporter une contribution significative aux organismes hôtes [Lili et al., 2010; Thomas, 2004]. Cet aspect est souvent nuancé par différents auteurs arguant qu'un comportement purement égoïste d'un plasmide, se maintenant au fil des générations par simple réplication, maintenance et transfert, ne semble pas viable sur le long terme d'un point de vue évolutif [Slater et al., 2008; Thomas, 2004]. Diverses stratégies ont été adoptées par les plasmides en diminuant leurs éventuels effets négatifs sur la croissance de l'hôte et/ou en contribuant à lui apporter un avantage évolutif significatif [de la Cueva-Méndez and Pimentel, 2007; Heuer et al., 2008]. Celles-ci incluent des adaptations de leurs mécanismes de réplication, ségrégation, d'incompatibilité, de maintenance, d'addiction génétique et de transfert afin d'optimiser leur stabilisation au sein du génome-hôte et leur transmission inter-génomique. Il existe de nombreux exemples où les plasmides contribuent fortement à l'adaptation de l'organisme bactérien hôte à son environnement en étant, par exemple, porteur de gènes de virulence ou de résistance à des composés tels que les antibiotiques. Enfin, en tant que particule extra-chromosomique auto-réplicative, chaque plasmide possède une dynamique de population spécifique qui dépend de ces mécanismes génomiques et de l'écologie de l'organisme dans lequel celui-ci est présent [Slater et al., 2008].

1.3 Architecture des réplicons : mécanismes structuraux

Il a longtemps été estimé que le nucléoïde bactérien était un amas compact d'ADN sans structure. Les chromosomes bactériens sont en fait organisés en plusieurs superenroulements alignés tels les perles d'un collier, cette structure dépendant de divers complexes protéiques et de petites molécules associées au nucléoïde [Thanbichler, 2010]. Les chromosomes bactériens sont orientés selon leur origine de réplication et le site de terminaison de la réplication, mais également en fonction de leur centromère *parS* (séquence de fixation des protéines ségrégatives). Ils possèdent de plus une orientation spatiale spécifique, les positionnements de leurs origine de réplication et site de terminaison dans la cellule dépendant rigoureusement du cycle cellulaire [Toro and Shapiro, 2010].

1.3.1 Protéines s'associant au nucléoïde

Les Nucleoid Associated Proteins (NAP) sont une classe de protéines se fixant à l'ADN et intervenant dans la structure spatiale du génome [Dillon and Dorman, 2010]. En agissant au niveau de la structure de l'ADN, ces molécules ont de plus un rôle régulateur de fonctions telles que la réplication, la recombinaison et la maintenance, mais aussi de la réparation et de la transcription de l'ADN [Azam and Ishihama, 1999]. Les principales NAP sont décrites Table 1.2 et leur propriétés des mieux connues sont présentées Table 1.3. Parmi les NAP majoritairement isolées d'*E. coli* HU, IHF, H-NS, StpA et Fis sont les plus abondantes [Johnson et al., 2005]. CbpA, CbpB (aussi connue sous le nom de Rob) et Lrp sont représentées de façon plus minoritaire parmi les bactéries [Johnson et al., 2005]. TABLE 1.2: Principales protéines s'associant au nucléoïde (NAP).

- HU (Heat Unstable) est une protéine de type histone. Elle semble impliquée dans la compaction de l'ADN et est présente chez la plupart des bactéries. Des homologues HU semblent exister chez les Eucaryotes et Archées. HU est de plus impliquée dans la réplication, la ségrégation et la division cellulaire, en permettant notamment la stabilisation de la protéine DnaA sur *ori*. HU intéragit avec les superenroulements de l'ADN de façon non-spécifique, avec l'ADN simple brin ainsi qu'avec l'ARN [Johnson et al., 2005].
- **IHF** (Integration Host Factor) est un paralogue de HU et, de même, est en lien avec la condensation de l'ADN ainsi qu'avec la réplication, la ségrégation et la division cellulaire [Johnson et al., 2005]. IHF reconnaît spécifiquement des séquences de 30 à 35 paires de bases (pb). Elle est trouvée chez différentes espèces de Protéobactérie (bactéries Gram-négatif) mais n'est pas décrite chez les bactéries Gram-positif. IHF semble intervenir avec Fis (*cf.* ci-après) au niveau d'*ori* et joue un rôle dans la régulation de l'initiation de la réplication [Johnson et al., 2005].
- **H-NS** (Histone-like Nucleoid Structuring protein) est une protéine structurale s'attachant de façon non-spécifique préférentiellement sur des régions riches en A+T, retrouvée chez de nombreuses bactéries Gram-négatif. H-NS contraint les superenroulements *in vitro* et influence la structure de l'ADN localement [Johnson et al., 2005]. C'est un homologue de StpA (*cf.* ci-dessous).
- **Fis** (Factor for inversion stimulation) est une protéine structurale jouant différents rôles de régulation (de nombreuses fonctions telles que la réplication et la ségrégation) en s'attachant à des centaines de sites sur l'ADN d'*E. coli* [Browning et al., 2010]. Fis est capable de "tendre" l'ADN et, ainsi, de promouvoir ou d'inactiver différents promoteurs [Dillon and Dorman, 2010].
- Lrp (Leucine-responsive regulatory protein) influence la transcription de 10% des gènes d'*E. coli* en agissant soit comme activateur, soit comme répresseur. En plus d'intervenir dans la régulation des gènes impliqués dans le métabolisme, la virulence ou l'expression du pilus, Lrp influence la structure du nucléoïde [Browning et al., 2010]. Son implication dans le cycle cellulaire d'*E. coli* a été démontré [Corcoran and Dorman, 2009]. Des homologues de Lrp tels que AsnC ou PutR (révélés par des analyses de similarité de séquences avec HMMER; *cf.* Chapitre 4) sont retrouvés chez l'ensemble des bactéries.
- **CbpA** (Curved-DNA binding protein textbfA) est un homologue de la protéine DnaJ. CbpA a la capacité de se fixer à l'ADN et est liée au cycle cellulaire. Elle contribue à la croissance à basse température des bactéries et est nécessaire pour un déroulement normal de la division cellulaire. CbpA est décrite chez de nombreuses bactéries [Dillon and Dorman, 2010].
- CbpB (Curved-DNA binding protein B) ou Rob (Right origin binding) est une protéine se liant à l'ADN au niveau de sites d'attache localisés à proximité de l'*ori* et de différents promoteurs (dont celui de son gène) et semble de ce fait être un facteur de transcription [Azam and Ishihama, 1999].
- **Dps** (DNA protection during starvation protein) est une protéine du cycle cellulaire qui, chez *E. coli*, contribue au compactage de l'ADN et au passage de la phase exponentielle de croissance à la phase stationnaire [Azam and Ishihama, 1999].
StpA (Suppressor of td mutant protein A) est un homologue de séquence de H-NS et est de même capable de courber la molécule d'ADN [Azam and Ishihama, 1999]. StpA peut former un hétéro-complexe avec H-NS et possède différentes activités régulatrices, notamment au niveau de la croissance cellulaire [Dillon and Dorman, 2010].

Table 1.3	: Propriétés	des NAP	majoritaires	${\rm des}$	bactéries.
	Adapté de	[Dillon and	Dorman, 2010]		

			-			
\mathbf{P}^{a}	\mathbf{C}^b	\mathbf{P}^{c}	\mathbf{T}^d	Motif	\mathbf{M}^{e}	\mathbf{Pr}^{f}
HU	oui	ND	oui	joint à l'ADN double brin ou simple brin;	$\sim 9kDa$	hétérodimère
				preference pour les zones riches en A+1		$(HU\alpha - HU\beta)$
IHF	ND	ND	Oui	(A/T)ATCAANNNNTT(A/G)	$\sim 11 kDa$	hétérodimère
						$(IHF\alpha - IHF\beta)$
H-	ND	Oui	ND	zone riche en A+T et (TCGATAAATT)	$\sim 15 kDa$	homodimère
NS						ou hátán dina ha
						(H-NS-StpA)
	<u> </u>	<u> </u>	<u> </u>		111 D	
Fis	Oui	Oui	Oui	zones riches en A ou A+T	$\sim 11 kDa$	homodimère
Lrp	Oui	Oui	ND	$(\mathrm{T/C})\mathrm{AG}(\mathrm{A/T/C})\mathrm{A}(\mathrm{A/T})\mathrm{ATT}(\mathrm{A/T})\mathrm{T}(\mathrm{A/T/G})$	$\sim 18 kDa$	homodimère
CbpA	ND	ND	ND	ADN courbe	$\sim 33 kDa$	monomère
Dps	ND	ND	ND	ND	$\sim 19 kDa$	monomère ou
						dodécamère
StpA	ND	Oui	ND	zone riche en A+T	$\sim 15 kDa$	homodimère
						ou
						heterodimere (StpA-H-NS)
14.15	NE	~ ·	ND			
MukB	ND	Oui	ND	ND	$\sim 175 kDa$	homodimere
Bactér	ries (Gram	posi	tif		
\mathbf{P}^{a}	\mathbf{C}^b	\mathbf{P}^{c}	\mathbf{T}^d	Motif	\mathbf{M}^{e}	\mathbf{Pr}^{f}
HU	ND	ND	Oui	ND	$\sim 10 k D a$	homodimère
Lrp	Oui	Oui	ND	ND	$\sim 17 kDa$	homodimère
MukB	ND	Oui	ND	préférence pour ADN simple brin	$\sim 130 kDa$	homodimère
^a P	: nom	de la N	JAP.			

Bactéries Gram négatif

 b C : capacité de la protéine à condenser l'ADN au niveau du motif de fixation.

 c P : capacité de la protéine à faire des "ponts" au niveau du motif de fixation.

 d T : capacité de la protéine à tendre l'ADN au niveau du motif de fixation

 e M : masse moléculaire en kilo Dalton.

 f Pr : promoteur natif.

Ces protéines, étant liées à la structure du nucléoïde, interviennent sur la transcription d'une manière globale. Les effets sur la transcription ne proviennent pas uniquement des changements du nombre de protéines disponibles au cours des différentes étapes du cycle cellulaire mais dépendent aussi du passage de la fourche de réplication [Browning et al., 2010]. Il existe de plus de nombreuses preuves d'interaction et de régulation entre les NAP aux niveaux géniques et protéiques. Par exemple, il existe une régulation croisée entre IHF, Fis et H-NS au niveau du promoteur de dps. H-NS inhibe les promoteurs de stpA. hns est lui même sous le contrôle de Fis. Lrp auto-régule négativement son propre gène et active stpA [Dillon and Dorman, 2010]. Enfin, il est à noter qu'il existe une interaction entre IHF et le super-régulateur CtrA (chez les alphaprotéobactéries) régulant l'initiation de la réplication en changeant l'architecture d'*ori* [Thanbichler, 2010].

1.3.2 Motifs structurels du nucleoïde

La régulation des mécanismes du cycle cellulaire implique dans presque tous les cas connus des intéractions entre éléments cis (séquences d'ADN) et trans (éléments diffusibles : protéines ou ARN régulateurs). Le cycle cellulaire est donc en quelque sorte reflété dans la séquence du nucléoïde. Les motifs, courtes séquences caractéristiques d'ADN, sont des cibles privilégiées des facteurs de transcription et de structure (*i.e.*, les NAP). Un inventaire non-exhaustif des principaux motifs impliqués dans la réplication, la ségrégation et la maintenance du chromosome bactérien [Touzain et al., 2011] est présenté Tables 1.4 et 1.5.

TABLE 1.4: Motifs structurels du nucleoïde.

Boîtes	Ces motifs sont impliqués dans le contrôle de l'initiation de la réplication en permet-
DnaA	tant la stabilisation de la protéine initiatrice DnaA ($cf.$ ci-après). Ces motifs (comme DnaA) sont conservés parmi les bactéries. Ils sont présents à l' ori du chromosome ainsi que sur des sites secondaires impliqués dans la séquestration de DnaA. Il existe différents types de boîtes DnaA sur lesquelles se fixent DnaA avec plus ou moins d'affinité selon sa conformation [Mott and Berger, 2007].
chi	Ces motifs (octamériques chez <i>E. coli</i>) interagissent avec les recombinases homologues RecBCD qui interviennent dans la réparation de l'ADN. Ces sites sont orientés et permettent de "guider" les recombinases [Spies and Kowalczykowski, 2005]. Ces sites sont très conservés dans toutes les espèces examinées [Touzain et al., 2011].
dif	Ce site, localisé à proximité du site de l'arrêt de la réplication des chromosomes bactériens, permet la fixation et l'action des systèmes de résolution de dimère de réplicon (systèmes Xer). Ces sites de type palindrome sont très conservés parmi les chromosomes bactériens et sont aussi trouvés chez les Archées [Carnoy and Roten, 2009].

- **GATC** Ces motifs de quatre nucléotides interagissent principalement avec Dam et SeqA, deux protéines impliquées dans l'initiation de la réplication ainsi que MutH. Selon leur état méthylé ou non, ils interviennent dans l'identification du brin ADN néo-formé dans les processus de réplication et de réparation de l'ADN. Ils semblent également avoir un rôle dans la ségrégation des chromosomes et sont des séquence cis-régulatrices de gènes. Ces motifs sont stratégiquement localisés par rapport à l'*ori*. La reconnaissance des GATC par Dam semble cependant être limitée aux gamma-protéobactéries [Touzain et al., 2011].
- KOPS (FtsK-Orienting Polar Sequences) Ces motifs orientés dans le sens de la réplication et donc inversés au site de terminaison de la réplication[Kono et al., 2011], guident les processus de résolution de dimère au site dif [Bigot et al., 2005]. Ils interviennent dans le guidage des recombinases site-spécifiques XerCD, responsables de la résolution de dimère de chromosomes au site dif avec la protéine FtsK, elle-même guidée par les motifs KOPS qui permettent sa fixation selon une orientation donnée.
- NBS (Noc-Binding Sites). Ces motifs de 14 pb chez *B. subtilis* sont les sites de fixation de la protéine Noc (Nucleoid occlusion protein) qui protège le chromosome d'une section par une mauvaise position du septum. Ils sont impliqués dans la prévention de l'assemblage de l'appareil moléculaire lié à la division cellulaire. SlmA est un analogue de Noc chez *E. coli* mais présente une séquence différente. Il est donc probable que, tout comme Noc, SlmA se fixe sur des sites particuliers.
- oriT Site de coupure des relaxases sur les éléments conjugatifs (§1.7.4). Ces sites peuvent être classés en différents groupes. Un même groupe peut rassembler des sites provenant de plasmides de bactéries Gram-négatif ou Gram-positif, ce qui suggère une origine commune à la conjugaison des éléments génétiques [Lawley et al., 2004].
- parS Ce site est l'homologue bactérien du centromère. Il stimule la ségrégation des chromosomes néo-formés par l'interaction avec ParA et ParB, des analogues du cytosquelette eucaryotique [Livny et al., 2007]. Les motifs parS sont des palindromes de 16 pb (chez Bacillus subtilis) et sont situés, en un ou deux exemplaires, à proximité de l'ori. Ceux-ci, tout comme le système ParA/ParB, sont très conservés dans le domaine bactérien à l'exception de certaines gamma-protéobactéries (Enterobactéries, dont E. coli) qui, elles, utilisent similairement le site migS [Livny et al., 2007; Mierzejewska and Jagura-Burdzy, 2012].
- ram (RacA binding motifs) Ces sites sont impliqués dans l'accrochage de l'ori aux pôles de la cellule via RacA lors de la sporulation chez B. subtilis.
- ter(pour terminaison) Ces sites font 23 pb de long et sont localisés dans la région opposée
à l'ori. Ils sont impliqués dans la terminaison de la réplication. Chez E. coli, ce sont les
sites de fixation de la protéine Tus qui intervient dans la déstabilisation des fourches
de réplication et l'arrêt de la réplication. Ils sont faiblement conservés parmi les
génomes bactériens.

Nom	Consensus	Nombre	Protéine	Fonction
Boîte	TTATNCACA (E. coli)	107	Dred	Initiation de la
DnaA	TTATNCACA (B. subtilis)	211	DhaA	réplication
GATC	GATC (E. coli)	19.120	Dam, SeqA, MutH	Identification du brin néo-synthétisé
	GN(A/G)NGTTGTAA(C/T)(T/G)A (E. coli)	10		
ter	$\begin{array}{c} ({\rm G/T})({\rm A/C}){\rm ACT}({\rm A/G}){\rm AN}({\rm A/T})({\rm A/G}) \\ ({\rm A/C/T})({\rm A/T})({\rm T/C})({\rm T/A})({\rm A/G}){\rm T} \ {\rm et} \\ ({\rm T/A})({\rm A/G}){\rm TG}({\rm T/A}){\rm AC}({\rm C/T}){\rm AA}{\rm A} \\ {\rm T}({\rm G/A/T}){\rm TT}({\rm C/T}) \ (B. \ subtilis) \end{array}$		Tus/ Rtp	Terminaison de la réplication
	GCTGGTGG (E. coli)	1.008		Réparation des cassures double-brin
Chi	GNTGGWGG (Haemophilus influenzae)	408	RecBCD/	
	AGCGG (B. subtilis)	11.381	AddAB	
	GAAGGGG (Staphylococcus aureus)	339	1	
	GCGCGTG (Lactococcus lactis	187		
parS	(T/C)GTT(T/A/C)CA(C/T)(G/A) TG(A/G/T)AAC(A/G) (B. subtilis)	10	SpoOJ	Ségrégation d'ori
migS	ATTTTTTGCGGGGTACTCAGCAAAATT (E. coli)	⁻ 1	inconnue	Ségrégation d'ori
ram	TGNCGCCGGCGNCA (B. subtilis)	923	RacA	Ségrégation d' <i>ori</i> lors de la sporulation
KOPS/ SRS	GGGNAGGG (E. coli)	366	FtsK	Ségrégation du chromosome

 TABLE 1.5: Caractéristiques des principaux motifs trouvés chez les bactéries.

 Adapté de [Touzain et al., 2011].

1.3.2.1 NAP et motifs structurels chez les plasmides

Tout comme les chromosomes, les plasmides possèdent une structure dont une partie est contrainte et maintenue par différents mécanismes [Higgins and Vologodskii, 2004]. Le spectre d'hôtes des plasmides est lié à leurs interactions avec les protéines structurales des hôtes (notamment IHF et Fis) [Del Solar et al., 1996; Toukdarian, 2004]. De nombreux NAP et les mécanismes impliqués dans la structure et la régulation des chromosomes ont donc également un rôle dans la structure des plasmides. Par exemple, IciA, Fis, IHF, HU interviennent dans la régulation de l'initiation de la réplication de certains plasmides [Del Solar et al., 1998; Krüger and Rakowski, 2004]. Fis se fixe à certains plasmides [Rimsky and Travers, 2011] et différents plasmides transmissibles codent pour un homologue de H-NS [Browning et al., 2010]. Des boîtes DnaA ont été identifiées au niveau de l'ori de différents plasmides $(cf. \S1.4.1)$ et des homologies de séquence existent entre les séquences centromériques plasmidiques et chromosomiques [Livny et al., 2007; Mierzejewska and Jagura-Burdzy, 2012]. Les plasmides semblent cependant posséder certaines spécificités structurales, telles que des origines de réplication organisées en itérons (motifs spécifiques permettant l'attachement des protéines initiatrices Rep) et des homologues des NAP chromosomiques et/ou de motifs structurels spécifiques.

1.4 Réplication des réplicons

La réplication chez les Eucaryotes, Archées et Bactéries présentent des similitudes : elle débute avec la fixation au niveau d'une région spécifique de l'ADN (origine) de la protéine initiatrice, ce qui provoque l'ouverture de la molécule double-brin et l'établissement d'un complexe de réplication. La réplication des chromosomes doit s'effectuer une seule fois durant le cycle cellulaire sous peine de létalité pour la cellule. La réplication est suivie d'une phase de ségrégation du matériel génomique néo-formé et de partition dans les deux cellules filles, chacune récupérant un complément génomique (Figure 1.2). Ces étapes sont sous le contrôle de différents mécanismes de maintenance coordonnant leurs déclenchement et inhibition en rapport avec le cycle cellulaire [O'Sullivan, 2011; Thanbichler, 2010]. La duplication des plasmides suit les mêmes processus, ceux-ci pouvant cependant posséder des mécanismes de régulation spécifiques.



FIGURE 1.2: Représentation schématique de la réplication du chromosome bactérien. Vert : *ori*. Rouge : sites de terminaison. Violet : réplisome. Adapté de [Ghosh et al., 2006].

Dans la réplication de tout réplicon, trois phases peuvent être distinguées :

- L'initiation qui correspond à l'ouverture de l'ori par des protéines initiatrices.
- L'élongation ou assemblage du brin néo-formé par les ADN-polymérases.
- La terminaison qui est la relaxation des complexes de réplication pouvant s'accompagner de la résolution de dimère.

La réplication chez les chromosomes bactériens est conservée et suit globalement le modèle décrit chez *E. coli*, alors qu'il existe trois types de réplication dans le cas des plasmides [Del Solar et al., 1998] :

- la réplication thêta, similaire à la réplication chromosomique,
- la réplication par rotation de cercle (RC),

• la réplication par déplacement de brin (SD).

Parmi les plasmides à réplication de type thêta, on peut différencier les plasmides à itérons, sites spécifiques d'interaction avec les protéines Rep au niveau de leur *ori*, et les plasmides de type "DnaA-like" [Petersen et al., 2011], des autres plasmides.

Les mécanismes et protéines majeurs impliqués dans la réplication, la ségrégation et le cycle cellulaire du génome bactérien sont résumés ci-dessous. La littérature existante fait majoritairement référence à quelques organismes modèles seulement, les principaux étant *E. coli* (gamma-protéobactérie), *Caulobacter crescentus* (alpha-protéobactérie) et *Bacillus subtilis* (firmicute).

1.4.1 Origine et initiation de la réplication

Toutes les cellules des trois domaines de la vie possèdent des mécanismes de régulation contrôlant la réplication durant la phase d'initiation [Clark and Maaløe, 1967]. Les protéines majeures impliquées dans l'initiation de la réplication des réplicons sont présentées Table 1.6. Sont ensuite détaillés la structure des origines de réplication, le déroulement et la régulation de l'initiation.

TABLE 1.6: Principales protéines impliquées dans l'initiation de la réplication.

DnaA	protéine s'attachant à l'ADN de façon spécifique. C'est un acteur principal de la cascade d'événements initiant la réplication des chromosomes bactériens et de cer- tains plasmides [Petersen et al., 2011]. Ces événements incluent la reconnaissance de l'origine (<i>ori</i>), son ouverture et l'attachement d'hélicases [Higgins, 2005]. DnaA est aussi un facteur de transcription [Messer, 2002]. On estime que des homologues de DnaA existent chez toutes les espèces bactériennes [Messer, 2002; Yoshikawa and Ogasawara, 1991].
DnaB	l'hélicase réplicative chez les bactéries. Elle est composée de six sous-unités et ca- talyse la séparation des brins de l'ADN [O'Donnell et al., 2013]. De même que pour DnaA, on considère que chaque organisme bactérien dispose d'une hélicase réplicative. Elle est de plus retrouvée dans les trois domaines du vivant.
DnaC	paralogue de DnaA. C'est un facteur supplémentaire permettant l'attachement et la stabilisation de DnaB à <i>ori</i> [Mott and Berger, 2007].
IciA	(Inhibitor of chromosome initiation) membre de la famille des facteurs de trans- cription LysR, très répandu chez les bactéries Gram négatif. IciA agit en tant qu'an- tagoniste de DnaA en inhibant l'initiation de la réplication au niveau d' <i>ori</i> [Dillon and Dorman, 2010].
DnaG	primase qui synthétise de courts fragments d'ARN, ou brins d'Okazaki, servant de points d'initiation pour la synthèse d'ADN [O'Donnell et al., 2013].
Dam	(D NA a denine m ethylase) méthyle les motifs GATC au niveau de l'adénine et par- ticipe à la régulation de l'initiation.

- **SeqA** (Sequestration A protein) régule l'initiation de la réplication en se fixant aux motifs GATC hémiméthylés au niveau d'*ori*. L'origine néo-formée est ainsi séquestrée, empêchant un nouvel amorçage de la réplication par DnaA.
- Hda Homologue de DnaA. Elle intervient dans la régulation de DnaA chez *E. coli*. Elle interagit avec l'ADN polymérase III, ce qui stimule l'activité de DnaA et contribue à l'ouverture de l'ADN au niveau d'*ori*. Des orthologues de Hda ont été trouvés uniquement chez les gamma-protéobactéries [Zakrzewska-Czerwińska et al., 2007]. Chez *B. subtilis*, c'est la protéine YabA qui joue ce rôle [Mott and Berger, 2007].
- SSB (Single Strand Binding proteins) En se fixant sur l'ADN mono-brin généré par l'ouverture de l'ori, ces protéines stabilisent le complexe d'initiation [O'Donnell et al., 2013].
- **DiaA** (DNA initiator-associating factor) protéine découverte relativement récemment chez *E. coli.* Elle se fixe à DnaA et est impliquée dans la synchronisation de l'initiation [Katayama et al., 2010].
- Rep Cette dénomination fait référence de façon générique aux protéines impliquées dans l'initiation de la réplication des plasmides. Elles sont apparentées par leur séquence aux protéines impliquées dans la conjugaison (Tra et Mob) [Del Solar et al., 1998] et sont retrouvées sur la plupart des plasmides dont la séquence complète a été réalisée. Elles peuvent comporter des domaines fonctionnels similaires [Del Solar et al., 1998], suggérant une origine commune. Tout comme DnaA avec le chromosome, les protéines Rep intéragissent avec les ori plasmidiques par la reconnaissance de motifs spécifiques. La désignation "Rep" englobe les protéines initiatrices nécessaires pour les trois types de réplication plasmidique bien qu'elles puissent correspondre à des protéines différentes avec des fonctions distinctes. Généralement les plasmides de type thêta codent pour un seul initiateur protéique, RepA, TrfA ou RepE, qui active ou inhibe la réplication selon sa configuration (dimère ou monomère) [Krüger and Rakowski, 2004]. Les plasmides de type SD codent pour trois protéines Rep : RepA (hélicase), RepB (primase) et RepC (facteur stabilisateur de l'hélicase de type DnaC) [Del Solar et al., 1998]. Les Rep des plasmides de type RC sont très conservées, et contiennent deux domaines spécifiques : le domaine dso de reconnaissance de l'origine et le domaine de coupure [Khan, 2005]. Ces protéines possèdent un résidu tyrosine qui est impliqué dans la coupure de l'ADN [Khan, 2005]. Elles recrutent une hélicase de l'hôte, PcrA, qui possède des homologies avec l'hélicase Rep d'*E. coli*.
- **DivK** (cell division response regulator **K**) régulateur antagoniste de CtrA lorsque il est activé [Brilli et al., 2010].
- YabA inhibiteur de DnaA chez *B. subtilis.* YabA forme un complexe avec DnaA et l'ADN polymerase III (*cf.* ci-après) et limite le nombre de protéines initiatrices disponibles [Katayama et al., 2010]. YabA est retrouvée chez d'autres bactéries Gram-positif [Mott and Berger, 2007].

SirA et (Sporulation inhibitor of replication protein A et Sporulation stage 0 protein A, respectivement) protéines impliquées dans la séquestration d'ori chez B. subtilis et ses proches voisins taxonomiques seulement.

1.4.1.1 Structure des origines de réplication

L'origine de réplication d'un chromosome bactérien typique est une courte séquence nucléotidique (250 pb chez *E.coli*; Figure 1.3), organisée en une région riche en Adénine et Thymine et en différents motifs structuraux [Rajewska et al., 2012; Robinson and Bell, 2005].



FIGURE 1.3: Origine de réplication de E. coli. Adapté de [Mott and Berger, 2007].

DUE : DNA Unwinding Element, riche en bases A+T. R1, R2, R3, R4 : boîtes DnaA orientées. Motifs présentant une affinité forte pour DnaA : bleu sombre, et faible : bleu clair. Rectangles blancs : sites de fixation de IHF et Fis. Étoiles : motifs GATC. Rectangles gris : sites I de fixation préférentielle de DnaA-ATP. Flèches : quatrième classe de motifs de fixation de DnaA-ATP exclusivement, dans DUE.

Sa proximité avec certains gènes engendre plusieurs systèmes de régulation. La plus importante interaction est avec DnaA, au niveau des boîtes DnaA qui structurent l'origine de réplication par leur type et leur orientation et permettent l'ouverture d'*ori* au cours d'un changement de conformation lors de la fixation de DnaA [Mott and Berger, 2007]. La région riche en bases A et T (DNA Unwinding Element; DUE) est le site d'ouverture proprement dit. Relativement moins d'énergie sera nécessaire à son ouverture de par sa richesse en paires A-T (2 liaisons covalentes) en comparaison à une région d'ADN comprenant des appariements G-C (3 liaisons covalentes). Cette région contient de plus, sur le chromosome d'*E. coli*, trois 13-mères caractéristiques organisés en tandem, impliqués dans l'activité de DnaA [Mott and Berger, 2007]. D'autres motifs structurels de cette région participent à la régulation de l'initiation, tels que les motifs d'attache de Fis et IHF et les motifs GATC. L'organisation des *ori* des chromosomes d'autres bactéries présentent de nombreuses similarités avec l'*ori* d'*E.coli* et respectent la structure DUE + motifs spécifiques (boîtes DnaA, site de méthylation de Dam, etc) (Figure 1.4).

La structure des ori plasmidiques présente certaines spécificités :

La structure des ori des plasmides à réplication thêta présente une organisation similaire à celles des chromosomes : DUE + motifs structuraux (Figure 1.4). De nombreuses origines plasmidiques comportent des boîtes DnaA [Rajewska et al., 2012], ce qui indique l'existence d'une interaction forte entre des régulateurs de l'hôte et la réplication plasmidique [Krüger and Rakowski, 2004]. Les ori plasmidiques peuvent également posséder des motifs spécifiques organisés en tandem, les itérons, qui permettent la fixation des protéines Rep et ainsi l'ouverture d'ori.



FIGURE 1.4: Diversité des origines de réplication des réplicons bactériens chromosomiques et plasmidiques. Adapté de [Rajewska et al., 2012].

- L'origine de réplication chez les plasmides de type RC est un peu particulière car elle comporte deux origines : dso (double strand origin) et sso (single strand origin). La première est le lieu d'ouverture de l'ADN double brin. Elle comprend deux sites spécifiques : nick, lieu de la coupure de l'ADN, et bind, lieu de fixation de la protéine Rep. C'est à son niveau que s'amorce la réplication d'un des deux brins. L'autre origine, sso, est l'endroit où est initiée la réplication de l'autre brin [Khan, 2005].
- Les plasmides SD sont majoritairement les plasmides de la famille IncQ [Loftie-Eaton and Rawlings, 2012]. Leur origine est structurée en trois itérons correspondant au site de fixation de RepC, deux régions riches en G+C et A+T, respectivement, et deux sites d'initiation simple brin : *ssiA* et *ssiB* (single strand initiation A et B).

1.4.1.2 Déroulement de l'initiation

 Dans le modèle classique de l'initiation de la réplication du chromosome, l'attachement de plusieurs DnaA sur les boîtes DnaA d'ori provoque une distorsion de l'ADN et conduit à son ouverture. Ce phénomène est suivi de la formation du complexe d'initiation (orisome), et de l'action de l'hélicase sur l'ADN ouvert [Robinson and Bell, 2005]. Chez E.coli, Fis initialement présent est déplacé, ce qui permet la fixation de IHF et la formation d'un complexe protéique : le *réplisome* [Mott and Berger, 2007; Zakrzewska-Czerwińska et al., 2007]. L'ouverture est stabilisée par la fixation de protéines SSB qui protègent l'ADN simple brin.

- L'initiation de la réplication thêta des plasmides est assez similaire à celle du chromosome bactérien : la fixation des protéines Rep sur les itérons provoque un changement de conformation de l'origine ce qui entraîne son ouverture. Des facteurs additionnels (DnaA, IHF, Fis, IciA, SeQ) peuvent intéragir avec Rep ou se fixer à l'ori afin de stabiliser/déstabiliser le complexe [Krüger and Rakowski, 2004].
- Chez les plasmides RC l'ouverture de l'ADN s'effectue au niveau du site nick du site sdo par fixation de la protéine Rep. Rep recrute ensuite une hélicase spécifique de l'hôte, PcrA (Plasmid copy number reduction protein A), qui étend l'ouverture. Un manteau de protéines SSB stabilise un des brins de l'ADN pendant que le complémentaire de l'autre brin est synthétisé par l'ADN polymérase III de l'hôte. La réplication du deuxième brin commence au site sso et fait également intervenir la machinerie réplicative de l'hôte : ADN polymérase I et III, ADN gyrase et ADN ligase [Khan, 2005].
- L'initiation de la réplication des plasmides SD commence par la fixation de RepC au niveau des itérons, ce qui induit une rupture de la structure de l'origine et conduit à son ouverture. RepA pénètre alors dans l'ouverture et catalyse le déroulement de l'ADN. Lorsqu'ils sont à découvert, les sites *ssiA* et *ssiB* sont reconnus par RepB qui permet la fixation de l'ADN polymérase III et le début de la réplication grâce aux amorces synthétisées par RepB [Loftie-Eaton and Rawlings, 2012].

Contrairement à certains plasmides à réplication thêta, les plasmides SD et RC nécessitent une intervention réduite des protéines de l'hôte. Il est alors intéressant de constater qu'au final, l'initiation de la réplication du chromosome n'est qu'un cas particulier de celle des plasmides thêta.

1.4.1.3 Régulation de l'initiation

Les modèles classiques de régulation de l'initation de la réplication impliquent majoritairement deux types de processus selon leur cible :

- La séquestration, la déstabilisation ou l'occupation d'ori et de ses motifs structuraux.
- L'inactivation, la séquestration ou la régulation des protéines initiatrices, Rep et DnaA.

L'initiation de la réplication du chromosome est très précisément régulée pour garantir qu'une unique réplication du chromosome aura lieu au cours d'un cycle cellulaire.

• Les mécanismes de séquestration de l'origine de réplication du chromosome chez *E. coli* font intervenir les motifs GATC (GANTC chez *B. subtilis*) et les protéines Dam qui méthylent ces sites au niveau de leur adénine. Avant initiation de la réplication, les sites GATC de l'origine sont doublement méthylés sur les brins direct et indirect (GATC étant un palindrome, si on considère son complémentaire). Juste après l'initiation, ces sites se retrouvent hémiméthylés car le brin néo-synthétisé n'a

pas encore été soumis à l'action de Dam. Cette hémiméthylation conduit à la fixation de SeqA, ce qui empêche la fixation de DnaA au niveau de la nouvelle origine [Krüger and Rakowski, 2004; Mott and Berger, 2007]. La séquestration de l'*ori* par SeqA et Dam semble être une caractéristique des Entérobactéries, mais divers mécanismes de séquestration existent pour d'autres bactéries [Zakrzewska-Czerwińska et al., 2007]. Chez *B. subtilis*, la séquestration d'*ori* se fait par deux protéines non homologues de SeqA, SpoOA et SirA [Katayama et al., 2010].

- La régulation de DnaA chez *E.coli* implique aussi la titration des protéines au niveau d'une séquence caractéristique, le locus *datA*, avec pour effet de réduire le nombre de DnaA actives disponibles.
- Le mécanisme RIDA (Regulatory inactivation of DnaA) [Mott and Berger, 2007] permet le passage de la forme active de DnaA, DnaA-ATP, à la forme inactive, DnaA-ADP, par l'intervention des protéines HdA et d'une des sous-unités de l'ADN polymérase III chez *E. coli* [Katayama et al., 2010]. Ces mécanismes semblent aussi exister chez d'autres bactéries [Zakrzewska-Czerwińska et al., 2007].
- DnaA a aussi un rôle de facteur de transcription en se fixant sur les promoteurs de certains gènes. Elle exerce aussi une auto-inhibition de sa synthèse en se fixant au promoteur de son gène [Zakrzewska-Czerwińska et al., 2007].

Différents mécanismes de régulation de l'initation sont spécifiques des plasmides.

- Des mécanismes de séquestration d'ori existent également chez les plasmides. Un exemple en est le handcuffing de l'origine de réplication des plasmides à réplication thêta et à itérons. Les ori des deux plamides néo-formés se lient via les protéines Rep fixées, empêchant leur accès [Krüger and Rakowski, 2004].
- Le contrôle de la disponibilité en protéines Rep chez les plasmides dépend notamment de l'auto-régulation de la transcription des gènes rep. Il peut aussi passer par la structure-même de l'ori qui peut titrer Rep grâce à ses itérons ou des sites annexes [Krüger and Rakowski, 2004]. Certaines protéines initiatrices Rep, activatrices dans leur forme monomérique, peuvent former des dimères et alors devenir inhibitrices de la réplication [Cervantes-Rivera et al., 2011; Krüger and Rakowski, 2004]. L'équilibre thermodynamique étant plutôt en faveur des dimères, l'introduction d'un plasmide dans un hôte qui abrite déjà un plasmide utilisant la même protéine Rep aura tendance à ne pas pouvoir se répliquer, l'initiation étant directement bloquée par les dimères de protéines Rep des plasmides présents. Ainsi les protéines Rep interviennent aussi dans les mécanismes dit d'"incompatibilité plasmidique".
- Des inhibitions par intervention d'ARN anti-sens agissant au niveau de la régulation de l'initiation plasmidique ont été caractérisées [Brantl, 2004]. L'inhibition intervient majoritairement dans le blocage de la traduction de l'ARNm des protéines activatrices Rep (par exemple, chez le plasmide R1 d'*E. coli*) ou en se fixant au niveau de leurs cibles. Dans le cas des plasmides de type "RepABC", le gène de l'activateur de l'initiation RepC étant inclus dans un opéron regroupant *repA* et *repB* (homologues de *parA* et *parB*, respectivement), la synthèse de RepC peut être pseudo-auto-inhibée par RepA et/ou RepB [Pinto et al., 2012], ou inhibée par un

ARN antisens dont le gène est présent à l'intérieur même de l'opéron [Cervantes-Rivera et al., 2011]. Enfin, de nombreuses protéines de l'hôte (DnaA, IHF, Fis) peuvent intervenir dans la régulation des gènes *rep* ou participer à l'inactivation des protéines Rep, ainsi que dans l'activation/inactivation des *ori* plasmidiques [Krüger and Rakowski, 2004].

1.4.2 Élongation

Après la formation du replisome, l'ADN est répliqué bidirectionnellement chez les chromosomes et la majorité des plasmides à réplication thêta, et unidirectionnellement chez les autres, les plasmides à réplication RC et SD et certains plasmides à réplication thêta. Les protéines clé formant les complexes de réplication organisés aux deux fourches de réplication, sont les suivantes :

- une hélicase faisant passer l'ADN du stade double-brin au stade mono-brin,
- des molécules stabilisatrices de l'ADN mono-brin (SSB),
- des primases synthétisant des amorces ARN nécessaires à la formation des fragments d'Okazaki,
- des polymérases synthétisant l'ADN et/ou remplaçant l'ARN par de l'ADN,
- des systèmes de correction d'erreur,
- des ligases liant les différents fragments synthétisés.

Deux types de protéines additionnelles sont nécessaires au bon fonctionnement des polymérases : des protéines d'attache à l'ADN (pinces), qui forment généralement un anneau glissant autour de l'ADN, et des protéines permettant l'établissement des protéines pinces sur l'ADN. Les principales protéines impliquées dans l'élongation sont présentées Table 1.7. Une bonne revue des différents mécanismes moléculaires impliqués dans l'élongation peut être trouvée dans [Johnson et al., 2005].

holoenzyme, ADN polymérase III	L'ADN polymerase III est un complexe comprenant différentes protéines (10 chez <i>E. coli</i>), incluant les polymérases cœur, la pince glissante et cinq sous-unités responsables du chargement de la pince sur l'ADN. Ces protéines ont été plus particulièrement détaillées chez les bactéries Gram négatif (<i>E. coli</i> surtout) mais semblent être retrouvées chez l'ensemble des
Pol III core	bactéries ainsi que chez les Archées [O'Donnell et al., 2013]. Ces protéines interviennent dans les fonctions polymérase (DnaE) et 3'-5' exonucléase (DnaQ).
DnaN	Pince circulaire de l'ADN polymérase III.
Complexe d'attache de la pince PolA	Chez <i>E. coli</i> , cinq sous-unités sont responsables du chargement de la pince : DnaX (partie mobile), HolA (ouverture de la pince), HolB (partie fixe), HolC (protéine de transfert) et HolD (stabilisateur). ou ADN polymérase I. Enlève les amorces ARN par sa fonction exo- nucléase

TABLE 1.7: Principales protéines impliquées dans l'élongation.

DNA ligase I	Lie les fragments d'Okazaki en un seul fragment continu.
GyrA et GyrB	protéines de type topoisomérase. Elles ont pour rôle de relaxer les super- enroulements créés dans l'ADN lors de la réplication et par l'action des hélicases. Les topoisomérases peuvent être classées en deux catégories se- lon qu'elles agissent sur l'ADN double ou simple brin. GyrA et GyrB sont les deux sous-unités de la topoisomérase II d' <i>E. coli</i> et agissent sur l'ADN double brin.

Une particularité de la réplication des plasmides SD, est qu'en plus d'être unidirectionnelle, il n'y a pas de production de fragment d'Okazaki pendant celle-ci. Les réplications SD et RC ont de plus la particularité de générer, durant la réplication, un intermédiaire d'ADN simple brin. Celui-ci étant hautement instable, ces types de réplication sont limités aux petits réplicons.

1.4.3 Terminaison de la réplication et résolution de dimère de réplicon

1.4.3.1 Mécanismes de terminaison de la réplication

Trois modèles décrivent actuellement la terminaison de la réplication des réplicons bactériens [Kono et al., 2012] :

- le modèle *fork collision* (collision de fourches) où la réplication bidirectionnelle s'interrompt lorsque les réplisomes des deux fourches de réplication se rencontrent.
- le modèle *fork trap* (piégeage de fourche) impliquant des associations ADN/protéine de type *ter*/Tus qui entravent la progression de la polymérase à des sites spécifiques. Une protéine (Tus chez *E. coli*, RTP (Replication Terminaison Protein) chez *B. sub-tilis* [Kono et al., 2012]) se fixe sur l'ADN de façon site-spécifique et intéragit avec les protéines du réplisome afin de les déstabiliser [Johnson et al., 2005; Kono et al., 2012].
- le modèle *dif-stop*, contrairement aux deux précédents, implique une terminaison de la réplication à un site unique et précis.

Le modèle type ter/Tus a été caractérisé pour le chromosome d'*E. coli*. La région de la terminaison est bi-polarisée par une distribution de motifs ter, orientés selon le sens de la réplication, de part et d'autre du site dif. La protéine Tus n'est pas conservée dans le domaine bactérien et son homologue fonctionnel chez *B. subtilis*, RTP, diffère tant au niveau structurel que par sa séquence, ce qui suggère une évolution relativement récente du modèle fork trap [Kono et al., 2012]. Les chromosomes ne possédant pas d'homologues fonctionnels des protéines *Tus* (*e.g.*, Firmicutes), ainsi que les plasmides à réplication thêta, semblent posséder un modèle d'arrêt de la réplication de type fork collision [Kono et al., 2012]. Les plasmides se répliquant de façon unidirectionnelle n'utilisent pas d'appareil de terminaison ni ne suivent le modèle par collision de fourches.

Dans la terminaison de la réplication des plasmides de type RC, les protéines Rep semblent impliquées en promouvant une coupure de l'ADN au niveau du site d'ouverture *dso*, ce qui libère un réplicon néo-formé ainsi qu'un intermédiaire simple brin [Del Solar et al., 1998; Khan, 2005]. Chez les plasmides SD la réplication finit de même au site d'ouverture, la réplication étant monodirectionnelle [Loftie-Eaton and Rawlings, 2012]. Les principales protéines jouant un rôle dans la terminaison de la réplication sont présentées Table 1.8.

TABLE 1.8: Principales protéines impliquées dans la terminaison de la réplication.

ParC	Sous-unité A de la topoisomérase IV. Trouvée chez $E. \ coli$ et $B. \ subtilis$ [Barnes et al., 2003].
ParE	Sous-unité B de la topoisomérase IV. Trouvée chez $E. \ coli$ et $B. \ subtilis$ [Barnes et al., 2003].
FtsK/SpoIIIE/ Tra	FtsK (Filamenting temperate sensitive protein K; SpoIIIE chez B. sub- tilis) Translocase essentielle impliquée dans le cycle cellulaire et coordon- nant les dernières étapes de la cytokinèse chez E. coli [Graham et al., 2010] en intéragissant avec ParC et XerD [Barre et al., 2000]. Sa struc- ture héxamérique entoure l'ADN et interagit avec FtsZ , FtsQ , FtsL et FtsI , autres protéines impliquées dans le cycle cellulaire. Cette fa- mille de translocases est très conservée chez les bactéries (sauf les Cya- nobactéries) [Bigot et al., 2007]. Tra , translocase trouvée chez certains éléments mobiles et impliquée dans la conjugaison, appartient à la même famille protéique [Bigot et al., 2007].
Tus	Cette protéine se fixe sur les motifs ter et déstabilise les protéines du réplisome lorsque celui-ci est à proximité d'un complexe Tus- ter . Elle est présente chez $E. \ coli$ et quelques Entérobactéries proches. Ni Tus, ni ter ne sont conservés parmi les chromosomes bactériens.
Recombinases site-spécifiques	cf. ci-dessous.

Une des caractéristiques essentielles des sites de terminaison de la réplication chez les bactéries est qu'ils sont directement liés aux mécanismes moléculaires impliqués dans la résolution de dimère de réplicon.

1.4.3.2 Résolution de dimère au niveau du site de terminaison

Au cours de la réplication des réplicons chromosomiques ou plasmidiques, des dimères peuvent apparaître, où les deux intermédiaires de réplication ne forment qu'une seule molécule via des processus de recombinaison homologue [Johnson et al., 2005]. La formation d'un dimère altère non seulement la ségrégation active des chromosomes et plasmides, mais perturbe également la réplication. Un multimère ayant plusieurs origines de réplication aura tendance à être sur-répliqué, ce qui peut conduire à l'épuisement de l'espace et des ressources disponibles de l'hôte [Hallet et al., 2004]. Dans le cas des plasmides, il y a multimérisation du plasmide jusqu'à ce que le nombre de molécules disponibles lors de la division cellulaire ne soit pas suffisante pour assurer sa stabilité [Summers and Sherratt, 1984]. Ces phénomènes ne sont pas marginaux; chez *E. coli*, 10 à 15% des réplications entraînent la formation de dimère [Pérals et al., 2000].

La résolution de dimère de chromosomes ou de plasmides requiert l'S (intervention d'une

machinerie moléculaire dédiée impliquant des recombinases spécifiques, dont les principales représentantes sont le couple de protéines XerC/XerD. Ces recombinases agissent au niveau d'un site spécifique (site *dif* des chromosomes) situé au niveau de la région de la terminaison de la réplication. Pour les plasmides, il existe différentes familles de recombinases qui présentent des mécanismes distincts (détaillées plus loin). Les sites de recombinaison sont eux-aussi différents de ceux des chromosomes [Hallet et al., 2004]. Contrairement au processus de recombinaison homologue, les systèmes de résolution de dimère impliquent une seule étape de recombinaison site-spécifique, indépendamment du cycle cellulaire pour les plasmides [Hallet et al., 2004] ou faisant intervenir FtsK pour les chromosomes.

Ces systèmes dédiés sont bien conservés parmi les chromosomes bactériens ainsi que pour de nombreux plasmides [Thomas, 2004].

1.4.3.3 Décaténation des régions terminales par les topoisomérases

Un autre obstacle à la séparation réussie des deux réplicons néo-formés est l'enchevêtrement des deux régions terminales [Thanbichler, 2010]. Les mécanismes impliqués dans la décaténation des deux réplicons font aussi intervenir FtsK qui interagit directement avec ParC et ParE, les deux sous-unités de la topoisomérase IV, impliquée dans la résolution des super-enroulements. En stimulant l'interaction de ParC et ParE entre la fin de la réplication et la séparation des cellules, FtsK permet l'activation de la topoisomérase IV [Thanbichler, 2010].

1.5 Partition active des réplicons

Après qu'un réplicon a été correctement répliqué, chacun des deux exemplaires obtenus doit être ségrégé dans une des deux cellules fille en formation. Cette ségrégation peut être passive comme, par exemple, dans le cas des petits plasmides à taux de copie élevé [Ebersbach and Gerdes, 2005] où les réplicons sont transmis de façon aléatoire à l'une des cellules fille. Dans le cas des réplicons à faible taux de copie, de grande taille et/ou d'une relative importance pour la vie de la bactérie, des mécanismes moléculaires sont nécessaires pour assurer une ségrégation réussie et la stabilité du matériel génomique au cours des générations. Il existe de grandes similarités entre les différents systèmes moléculaires de **partition** permettant une ségrégation active, présents chez les réplicons bactériens [Funnell and Slavcev, 2004]. Les acteurs majoritaires de ces systèmes sont les protéines de type **ParA** (ParA-like), **ParB** (ParB-like), ainsi que la séquence de type centromérique **parS** (Table 1.9).

TABLE 1.9: Principales protéines impliquées dans la ségrégation des réplicons.

MRP protéine fixant l'ATP, impliquée dans la ségrégation des chromosomes.

ParA/ Par commodité ces protéines seront désignées sous l'appellation : ParA-like et les gènes les codant, parA-like. Les protéines ParA-like peuvent agir en tant SopA/ que répresseurs des gènes parA et/ou parB. Ce sont des NTPases qui utilisent ParM l'énergie libérée par hydrolyse de l'ATP et mettent en mouvement les réplicons par des phénomènes de polymérisation. Selon les ParA-like, différentes configurations peuvent être prises par les polymères, bien que la conformation générale semble être similaire à celle des microtubules. Trois catégories de ParA-like peuvent être distinguées : les Walker-A P-loop ATPases (ParA, SopA ; type I), les actin-like ATPases (ParM; type II) et les tubulin-like ATPases (TubZ; type III) [Funnell and Slavcev, 2004; Mierzejewska and Jagura-Burdzy, 2012]. Deux sous-types de ParA type I sont identifiés selon leur structure : la (ParA des plasmides F et P1) et lb (protéines plus courtes) [Passot et al., 2012]. Les ParA-like de type II (ParM) ont la capacité de former des polymères avec une architecture très similaire à celle de l'actine chez les eucaryotes. Les ParA-like de type III ont été caractérisées plus récemment chez des plasmides de Bacillus cereus [Zheng et al., 2013]. Ce sont des GTPase, analogues des tubulines des eucaryotes. Les ParA des chromosomes sont usuellement de type I à quelques exceptions près ($E. \ coli$ notamment). MreB, protéine de structure liée au cycle cellulaire et vraisemblablement remplaçant ParA chez E. coli, est proche de ParM structurellement [Mierzejewska and Jagura-Burdzy, 2012].

ParB/ SpoB/ ParR

Par commodité ces protéines seront désignées sous l'appellation ParB-like et les gènes les codant, parB-like. Dans les premières phases de la ségrégation des réplicons, ces protéines s'attachent à leur séquence centromérique respective, parS (parfois appelée parC), et promeuvent la polymérisation des ParA-like [Funnell and Slavcev, 2004]. Les ParB-like possèdent aussi un rôle dans la régulation des systèmes de partition. Surexprimées, ces protéines, dans certains cas, inhibent les gènes autour de parS [Funnell and Slavcev, 2004]. Les ParB-like peuvent aussi jouer un rôle dans l'accrochage des réplicons à la membrane bactérienne via d'autres protéines intermédiaires [Mierzejewska and Jagura-Burdzy, 2012; Toro and Shapiro, 2010].

Ce système de partition aurait un fonctionnement similaire à celui de l'appareil mitotique eucaryote [Mierzejewska and Jagura-Burdzy, 2012]. L'étape initiale de la ségrégation de deux réplicons néo-formés débute par la fixation de la protéine ParB-like sur la séquence centromérique parS. Une fois assemblé, ce complexe peut interagir avec la protéine ParAlike. Pour être actives, ces dernières nécessitent l'action de l'ATP. Différentes structures peuvent être formées et différentes modalités de ségrégations mises en œuvre selon la protéine ParA-like impliquée [Funnell and Slavcev, 2004; Mierzejewska and Jagura-Burdzy, 2012]. Le modèle général semble néanmoins impliquer une polymérisation des ParA-like, ce qui a pour effet de "pousser" les réplicons aux pôles opposés d'une manière analogue à celle des microtubules chez les Eucaryotes [Mierzejewska and Jagura-Burdzy, 2012]. La régulation des systèmes de partition se fait d'une part par la dépolymérisation naturelle des ParA-like, et d'autre part par une régulation des gènes parA et parB, généralement regroupés dans un même opéron. Cette régulation fait intervenir les protéines ParA-like et ParB-like (autorégulation) ou des protéines annexes liées au cycle cellulaire [Funnell and Slavcev, 2004; Mierzejewska and Jagura-Burdzy, 2012; Pinto et al., 2012]. Outre le système parABS, d'autres systèmes moléculaires peuvent ségréger

les chromosomes bactériens. MukB et SMC (deux analogues) semblent par exemple être impliqués dans la ségrégation des chromosomes de E.~coli et de B.~subtilis, respectivement [Toro and Shapiro, 2010]. Le système parABS est retrouvé chez la très grande majorité des réplicons bactériens, ce qui est en faveur d'une émergence ancienne et unique, potentiellement existant dès l'origine des premiers réplicons. Ces systèmes constituent de fait un point clé de caractérisation parmi les propriétés fondamentales des réplicons.

1.6 Maintenance des réplicons et intégration dans le cycle cellulaire

Il existe un lien logique des contrôles de la réplication et des systèmes de ségrégation des réplicons avec le cycle cellulaire. Les divers mécanismes de maintenance développés par les réplicons permettent d'optimiser, en terme de coût énergétique et d'efficacité (*i.e.*, minimiser le taux de processus défectueux), leur réplication/ségrégation au sein de l'hôte et leur maintien dans l'espace et le temps à travers le cycle cellulaire (Table 1.10). De nombreux mécanismes moléculaires impliqués dans la structure des réplicons interviennent de fait dans la maintenance des réplicons. Classiquement, la fin de la réplication des chromosomes doit être suivie d'un détachement réussi des régions terminales des deux chromosomes, de processus de partition/ségrégation et des phénomènes de condensation selon un ordre précis [Thanbichler, 2010].

 TABLE 1.10: Principales protéines impliquées dans la maintenance et le cycle cellulaire des réplicons bactériens.

AcrA	fait partie d'un complexe protéique servant au transfert de diverses molécules à travers la membrane. C'est un acteur clé dans le cycle cellulaire [Lau and Zgurskaya, 2005]. Son inhibition influence directement le fonctionnement moléculaire des protéines du divisiome [Li et al., 2011]
AmiC	est une protéine périplasmique clôturant la formation du divisiome [Vicente et al., 2006].
EzrA	est un régulateur négatif de l'assemblage de l'anneau de FtsZ chez $B.\ subtilis$ [Yamanaka et al., 1996].
Fic	(Filamentation induced by cyclic AMP protein) est une protéine inductrice du cycle cellulaire chez $E. coli$ [Kawamukai et al., 1989]. Elle est aussi présente chez $B.$ subtilis.
$\mathbf{Fts}\mathbf{Z}$	Protéine majeure impliquée dans la formation du septum et dans la division cellulaire. FtsZ s'assemble en anneau à l'équateur de la cellule, ce qui per- met, par un mécanisme de constriction, de couper la membrane [Thanbichler, 2010]. FtsZ, très conservée parmi les bactéries (à une exception : <i>Chlamydia</i> , [Li et al., 2011]), est également présente chez certaines Archées [Vicente et al., 2006].
FtsA	Protéine intervenant avec ZipA dans la formation de l'anneau FtsZ [Vicente et al., 2006].

ZipA	s'accroche à FtsZ <i>in vitro</i> et aide à la stabilisation de l'anneau de FtsZ <i>in vivo</i> [Li et al., 2011]. Tout comme FtsB, FtsL et FtsN, ZipA n'est pas retrouvée chez certaines bactéries [Li et al., 2011].
ZapA	Protéine participant avec FtsA et ZipA à la formation de l'anneau FtsZ [Vicente et al., 2006].
FtsB, FtsE, FtsI, FtsN, FtsQ, FtsW, FtsX	protéines intervenant, après FtsA/ZipA, dans la stabilisation de l'anneau FtsZ [Vicente et al., 2006].
GidA	(Glucose inhibited division protein A) protéine quasiment universelle dans les génomes bactériens, influe sur la réplication [Kinscherf and Willis, 2002]. Ces protéines possèdent un domaine fonctionnel de type méthyltransferase.
GidB	(Glucose inhibited division protein B) est une protéine possédant, comme GidA, un domaine fonctionnel de type méthyltransferase. Tout comme GidA, GidB est probablement impliquée dans le cycle cellulaire [Ogasawara and Yoshikawa, 1992]. On la trouve aussi bien chez <i>B. subtilis</i> que chez <i>E. coli</i> .
TrmFO	est une protéine de la même famille que GidA qui semble également impliquée dans la division cellulaire [Cicmil, 2008]. Elle est généralement trouvée chez les bactéries à Gram positif.
\mathbf{SepF}	est une protéine conservée chez les bactéries à Gram positif. Elle interagit avec FtsZ (chez $B.$ subtilis) et joue un rôle significatif dans le développement du septum [Hamoen et al., 2006].
\mathbf{SlmA}	Protéine fixant l'ADN et inhibitrice de l'assemblage de FtsZ [Thanbichler, 2010]. SlmA, présente chez <i>E. coli</i> , est très peu conservée parmi les bactéries. Noc est son homologue fonctionnel chez <i>B. subtilis</i> .
SulA	est une protéine du système SOS (réparation sur épreuve) caractérisée chez <i>E. coli.</i> Elle est aussi impliquée dans l'arrêt de la division cellulaire en inter- agissant avec FtsZ [Yamanaka et al., 1996].
MinC	est un inhibiteur de la formation de polymères de FtsZ par oscillations rapides entre les pôles de la cellule [Thanbichler, 2010].
MinD	agit en concert avec MinC en s'attachant à celle-ci afin d'empêcher la po- lymérisation de FtsZ [Thanbichler, 2010].
MinE	est organisée en anneau autour de la membrane et déplace progressivement la position des complexes MinCD [Thanbichler, 2010].
MreB	est homologue de l'actine et est essentiel dans le maintien de la forme de bâtonnet d' <i>E. coli.</i> Son inactivation entrave sérieusement la ségrégation des chromosomes [Ebersbach and Gerdes, 2005]. MreB s'associe avec MreC et MreD. Elle est trouvée chez l'ensemble des bactéries.

MreC et MreD	sont impliquées, en interaction avec MreB, dans la maintenance de la structure en forme de bâtonnet de la cellule et dans la ségrégation des chromosomes $E.\ coli\ [$ Wachi et al., 1989].
MukB	joue un rôle central dans la condensation et la ségrégation de l'ADN, et le cycle cellulaire d' <i>E. coli</i> [Thanbichler, 2010; Yamanaka et al., 1996]. Cette protéine forme un complexe avec MukE et MukF et interagit avec FtsZ et la topoisomérase ParC .
MukE et MukF	MukF interagit avec MukB. MukF et MukE interagissent ensemble [Yama-naka et al., 1996].
SMC	(Structural Maintenance of Chromosome) est un analogue de MukB. Elle joue un rôle central dans la condensation du chromosome et est liée à ScpA et ScpB.
ScpA et ScpB	sont deux protéines auxiliaires du complexe SMC. Elles ont été découvertes chez les bactéries à Gram positif et chez les Archées [Thanbichler, 2010]. ScpA s'accroche à SMC et est stabilisée par ScpB. Des mutations sur ScpA provoquent le même effet que des mutations sur SMC [Thanbichler, 2010].
DivIVA	caractérisée chez <i>B. subtilis</i> , est une protéine majeure du cycle cellulaire. Elle intervient en intéragissant ave minCD et régule la formation du septum [Edwards and Errington, 1997]. Elle est de plus associée à RacA [O'Sullivan, 2011].
RacA	est une protéine d'ancrage chez <i>B. subtilis.</i> Elle accroche <i>ori</i> aux pôles de la cellule afin de préparer la division cellulaire. Elle est associée à DivIVA [O'Sullivan, 2011].
RodA	est une protéine impliquée dans la division et l'élongation cellulaire. Elle est décrite chez <i>B. subtilis</i> et <i>E. coli</i> [Henriques et al., 1998].
Systèmes PSK	(Post Segregational Killing) Véritable système d'"addiction génique", ces complexes géniques interviennent de façon indirecte dans la stabilité et la maintenance d'un élément génétique (plasmide, séquence d'insertion, phage, transposon, complexe allélique). Ces systèmes sont classiquement organisés en un gène codant une toxine et un gène codant une antitoxine avec éventuellement un gène régulateur. Une modification dans le système (perte de l'antitoxine, perte du système, gradient de concentration) peut entrainer une concentration trop importante de la toxine et la mort cellulaire. Il existe une variété de systèmes PSK : higBA, mazEF, relBE, HOK/SOK, vapXD, parDE, epsilon-zeta, ccd, Phd/Doc [Kobayashi, 2004].

Les mécanismes moléculaires de la division cellulaire impliquent la protéine FtsZ comme "fer de lance". Chez *E. coli*, FtsZ forme une structure polymérique en forme d'anneau dirigeant le processus de la division cellulaire [Thanbichler, 2010]. L'intervention de protéines additionnelles est souvent requise pour contrôler la position et la formation du *divisiome* (système protéique impliqué dans la division) telles que le système Min (impliquant les protéines **MinC**, **MinD** et **MinE**) et **SlmA** (chez *E. coli*), Noc (chez *B. subtilis*) ou le système **MipZ** chez *Caulobacter crecentus* [Thanbichler, 2010]. Une quinzaine de protéines supplémentaires (notamment **FtsA**, **FtsB**, **FtsE**, **FtsI**, **FtsN**,

FtsQ, FtsW, FtsX, ZipA, ZapA et AmiC) particulièrement bien conservées parmi les réplicons bactériens, sont recrutées dans la formation du septum [Vicente et al., 2006]. Chez les plasmides, la réplication, bien qu'asynchrone par rapport au cycle cellulaire, doit être régulée pour i) éviter une multiplication anarchique des plasmides et aboutir à un épuisement des ressources de l'organisme ("runaway replication" [Del Solar et al., 1998]), ii) dupliquer le matériel génétique au bon moment au bon endroit et, de ce fait, adapter le cycle plasmidique au cycle cellulaire, et iii) être transmis équitablement aux deux cellules fille de l'hôte originel [Pinto et al., 2012]. Différents mécanismes moléculaires spécifiques sont utilisés par les plasmides pour se maintenir à moindre coût, en adéquation avec leur nature non-essentielle, potentiellement égoïste et asynchrone par rapport au cycle cellulaire. Les mécanismes dits d'addiction plasmidique, ou Post-Segregational Killing (PSK), très largement représentés chez les plamides, sont caractérisés par l'installation dans le génome d'un complexe génique impliqué dans la production d'une toxine et de son antitoxine correspondante. Toute modification du complexe provoque la mort de l'hôte [Mochizuki et al., 2006]. Ces systèmes sont avantageux pour les plasmides d'un point de vue évolutif car ils garantissent le maintien des plasmides dans l'hôte les hébergeant. Ils sont aussi présents sur les chromosomes et permettent, par exemple, la stabilisation d'un complexe allélique Mochizuki et al., 2006]. De facon générale, il est intéressant de constater que l'évolution des génomes a gardé de tels systèmes géniques au comportement égoïste : outre le maintien de populations plasmidiques Mochizuki et al., 2006], ces systèmes peuvent contribuer à la maintenance chromosomique, comme par leur implication dans la survie de populations bactériennes en situation de carence ou faisant face à des éléments génétiques intrusifs (phages, transposons...) [Kobayashi, 2004]. Ces systèmes sont également la preuve que l'évolution des génomes bactériens a adapté des systèmes plasmidiques dans les processus de maintenance chromosomique (ou inversement).

1.7 Mécanismes moléculaires de recombinaison, d'intégration et de transfert génétique

Les recombinaisons génétiques sont des phénomènes d'échange de fragments entre deux molécules d'ADN. Elles nécessitent l'action de diverses protéines et enzymes dont les recombinases. On distingue la recombinaison homologue où des brins d'ADN sont échangés entre molécules homologues, des recombinaisons dite "site-spécifiques" pour lesquelles les enzymes agissent seulement au niveau de sites présentant des motifs structuraux particuliers. Les recombinaisons homologues sont majoritairement impliquées dans les phénomènes de réparation de l'ADN chromosomique pendant la réplication, où un des brins d'une première molécule d'ADN (double-brin) va recombiner avec un brin d'une deuxième molécule d'ADN endommagée et servir de matrice à différentes enzymes afin de corriger la lésion [Barre and Sherratt, 2005; Perry and Staley, 1997]. Elles interviennent aussi dans différents processus moléculaires chez les phages [Lopes et al., 2010]. Contrainment aux processus de recembinaison homologues homologues avec methinaisons des phages site

Contrairement aux processus de recombinaison homologue, les recombinaisons sitespécifiques sont réalisées par une machinerie moléculaire relativement simple où les recombinases vont prendre en charge les étapes de césure, d'échange et de fermeture de l'ADN [Hallet et al., 2004]. Elles ont lieu sur de courts segments d'ADN, les *sites de recombinaison*. Les recombinaisons site-spécifiques interviennent notamment dans la résolution de dimère précédemment abordée, mais aussi, de façon très similaire, dans la réplication de certains transposons [Hallet et al., 2004]. Enfin, elles peuvent intervenir dans les processus d'intégration et d'excision des bactériophages et de différents éléments mobiles [Hallet et al., 2004].

1.7.1 Résolvases impliquées dans la résolution de dimère de réplicon

Plasmides et chromosomes, dans leur grande majorité, codent des systèmes protéiques de recombinaison homologue spécifiques à la résolution des dimères formés lors de la réplication. Les enzymes impliquées sont des **résolvases**. On en distingue deux familles : les tyrosine- et les sérine-recombinases (Table 1.11). Cette distinction repose sur la nature de l'acide aminé servant en premier à l'attaque nucléophile de l'ADN par la recombinase.

TABLE 1.11: Résolvases impliquées dans la résolution de dimère.

Tyrosinerecombinases

Les tyrosine-recombinases présentent différents niveaux de complexité. Leur fonctionnement ne nécessite qu'un simple site de recombinaison ou peut impliquer des sites d'attache de protéines régulatrices. On trouve parmi elles les résolvases de type Xer qui sont un parfait exemple de protéine multiusage, à mécanisme simple, et largement répandues parmi des plasmides et chromosomes bactériens. Le système Xer le mieux décrit fait intervenir deux protéines, XerC et XerD, qui effectuent la recombinaison homologue des réplicons au niveau du site dif. Des homologues de ces protéines sont trouvés dans le génome de presque toutes les bactéries et archées ayant un chromosome circulaire [Cortez et al., 2010; Hallet et al., 2004]) et chez de nombreux plasmides, ce qui suggère une conservation stricte du système Xer/dif pour les chromosomes bactériens [Carnoy and Roten, 2009; Kono et al., 2011]. Il existe cependant des variations mono-protéiques au système XerCD. Dans le génome des firmicutes, une seule protéine homologue à XerC et XerD, XerS, fait office de résolvase [Leroux et al., 2011]. Un autre système Xer chez une partie des epsilon-protéobactéries implique une unique recombinase, XerH, qui ne présente pas de lien phylogénique apparent avec XerS [Carnoy and Roten, 2009]. Le système Xer/dif n'est cependant pas présent de manière universelle. Certains génomes, principalement chez des endosymbiotes, ne semblent pas posséder de sites dif, ni de protéines Xer [Carnoy and Roten, 2009; Kono et al., 2011]. D'autres systèmes de recombinaisons, faisant notamment intervenir des sérines-recombinases, peuvent remplacer les systèmes Xer/dif [Carnoy and Roten, 2009]. Des résolvases différentes de la famille Xer peuvent exister et participer à la résolution des dimères chez certains plasmides, comme par exemple, les résolvases de la famille de ResD découvert chez un plasmide d'E. coli [Hallet et al., 2004]. Enfin, les systèmes Xer/dif chromosomiques fonctionnent différemment de ceux des plasmides. Ils requièrent l'intervention de protéines supplémentaires (FstK entre autres, cf. §1.7) les liant au cycle cellulaire.

Sérinerecombinases Elles sont présentes sur les plasmides des bactéries à Gram positif ou négatif. Elles sont homologues des résolvases impliquées dans l'insertion de transposons et en sont clairement dérivées, pouvant d'ailleurs se substituer à celles de ces derniers [Hallet et al., 2004]. Le site de recombinaison est généralement proche des gènes codant les recombinases, ce qui facilite le transfert latéral d'un système fonctionnel. On en distingue plusieurs sous-familles : les recombinases de type Mu (proches des protéines du bactériophage Mu), de la famille Tn3 et les bêta-recombinases codées par certains plasmides de bactéries à Gram positif [Hallet et al., 2004].

Les chromosomes bactériens circulaires codent des tyrosine-recombinases, dont la grande majorité appartient au système XerC/XerD. Des homologues de ces protéines ont été identifiés chez de nombreux plasmides ainsi que chez des archées à génome circulaire [Hallet et al., 2004]. Les génomes linéaires ont des systèmes protéiques particuliers impliqués dans la résolution des télomères. Ces enzymes, les **ResT**, possèdent peu d'homologie de séquence avec les tyrosine-recombinases [Chaconas and Chen, 2005]. Les systèmes de résolution de dimère semblent avoir des fonctions annexes dans de nombreuses situations. Par exemple, ils peuvent être détournés par les phages pour s'insérer dans le génome de l'hôte ou permettre la conjugaison de leur génome [Das et al., 2013; Hallet et al., 2004].

1.7.2 Les autres recombinases

Les familles des tyrosine- et sérine-recombinases incluent par ailleurs les "**transposases**", impliquées dans l'insertion/excision des transposons, les "**relaxases**", initiatrices de la réplication (Rep) et de la mobilisation (Mob) des plasmides RC, les "**intégrases**" qui catalysent les étapes d'intégration des rétrotransposons et rétrovirus, et certaines topoisomérases (Table 1.12). De manière générale, ces protéines catalysent des opérations d'excision/intégration de l'ADN et tiennent une place capitale dans les cycles des différents éléments génomiques mobiles (ou MGE pour *Mobile genetic element*).

TABLE 1.12: Les recombinases distinctes des résolvases.

DDE transposases	Ce sont des enzymes impliquées dans les cycles des transposons bactériens. L'action catalytique de ces enzymes impliquent trois acides aminés (Asp (D), Asp (D) et Glu (E), formant le motif DDE), qui coordonnent les trans- ferts de groupement phosphoryle lors de la transposition [Higgins, 2005]. Ces transposases sont aussi présentes sur certains rétrovirus et rétrotranspons d'eucaryotes.
Intégrases et transposases de type tyrosine- recombinase	Le système $\operatorname{Cre}/\operatorname{lox}P$ est un système de recombinaison trouvé chez le bactériophage P1 d' <i>E. coli.</i> Il implique une unique recombinase Cre proche de l'intégrase Int du bactériophage lambda. Ce système est impliqué dans la circularisation des phages [Hallet et al., 2004]. Les tyrosine-recombinases comportent de plus des transposases codées par des transposons : Y2 ou RCT (pour Rolling Circle Transposition) [Higgins, 2005], proches des protéines Rep des plasmides RC [Cornet and Chandler, 2004], et les transposons Y codant une transposase plus classique.

Intégrases et	Plusieurs sérine-recombinases sont présentes en tant que transposases
transposases	chez les transposons S [Higgins, 2005]. Tout comme pour les tyrosine-
de type sérine-	transposases, le site de transposition est moins contraint que le site
recombinase	de résolution [Cornet and Chandler, 2004]. Les intégrases des sérine- recombinases sont représentées par Hin et Gin, deux enzymes plasmidiques [Cornet and Chandler, 2004].
Relaxases	Cette famille hétéroclite d'enzymes regroupe en particulier les protéines qui se fixent sur le site de conjugaison des éléments conjugatifs et mobi- lisables, <i>oriT</i> , et procèdent à son ouverture [Guglielmini et al., 2013]. Ces protéines sont longues et comportent deux, ou plus, domaines fonctionnels (un domaine relaxase et un domaine hélicase). Différentes familles de re- laxases sont identifiables et peuvent servir à classer les éléments conjugatifs [Guglielmini et al., 2013]. Ces relaxases sont de plus structurellement homo- logues des protéines initiatrices de la réplication de type RC [Smillie et al., 2010].

1.7.3 Les éléments transposables

1

Les processus de recombinaison spécifique sont fortement impliqués dans les processus de ségrégation et de maintenance des réplicons bactériens. Ils sont à la base de la définition d'une classe d'éléments génétiques distincte des réplicons, les transposons (Table 1.13), bien que le type de recombinaison manifesté par les transposons diffère de la recombinaison site-spécifique impliquée dans la résolution de dimère de chromosome ou de plasmide. Différentes nomenclatures et types de transposons ont été définis selon leurs mécanismes d'insertion, les transposases impliquées, leurs sites d'insertion ou l'organisation de la séquence transposable [Higgins, 2005; Roberts et al., 2008].

TABLE 1.13: Principaux types de transposons.

Séquences d'insertion	(IS) ou transposons unitaires. Ces séquences sont uniquement composées des structures (site de recombinaison) et gènes (transposases) impliqués dans la transposition.
Transposons composites	sont composés d'IS et de gènes et structures additionnels codant, par exemple, des protéines de résistance aux antibiotiques.
Transposons conjugatifs	ou ICE (<i>Integrative Conjugative Element</i>) portent des gènes codant des fonctions d'excision et d'intégration (transposases/intégrases) et des fonctions de conjugaison [Wozniak and Waldor, 2010].
Transposons mobilisables	ou IME (pour <i>Integrative Mobilisable Element</i>) peuvent être mobilisés entre cellules bactériennes, et codent en général leurs propres transposases [Roberts et al., 2008].
Îlots génomiques mobiles	sont des séquences intégrées dans les génomes, codant leurs propres trans- posases, mais pas les gènes impliqués dans leur transfert inter-cellulaire. Ils contiennent des gènes modifiant le phénotype cellulaire (par exemple, îlots de pathogénicité) [Boyd et al., 2009].

35

Prophages	
${\it transposables}$	

sont des génomes de bacteriophage intégrés dans des génomes bactériens et capable d'excision/insertion.

Prophages satellites

sont des prophages déficients intégrés dans des génomes bactériens. Ils nécessitent l'intervention de la machinerie moléculaire d'autres prophages pour compléter leur cycle de réplication.

Ces éléments sont capables de s'insérer, en tant que fragments d'ADN discrets et non permutables, à différents sites dans les génomes [Higgins, 2005]. Par opposition aux réplicons, les éléments transposables ne sont pas capables de se maintenir et de se répliquer indépendamment de leur génome hôte. La transposition requiert l'action de transposases (cf. §1.12) qui catalysent les différentes réactions de coupure/jointure de l'ADN. Même si originellement les transposons ont été décrits comme codant des DDE transposases, une grande variété de transposons codent des transposases de type sérine-ou tyrosine-recombinase, dont certaines sont homologues des relaxases et des protéines Rep des réplicons RC [Roberts et al., 2008]. Les transposons peuvent être soumis à des mécanismes de régulation contrôlant leur cycle d'excision/intégration, évitant de potentiels effets délétères à leur hôte. Ces mécanismes peuvent faire intervenir des ARN anti-sens, des protéines de méthylation (*i.e.* Dam), des represseurs transcriptionnels ou des protéines de type NAP de l'hôte [Cornet and Chandler, 2004]. L'existence sur des ICE d'analogues de systèmes de partition a également été rapporté [Wozniak and Waldor, 2010].

1.7.4 Conjugaison des éléments génomiques

Un des derniers aspects de la description du matériel génomique bactérien concerne les mécanismes et structures impliqués dans la mobilisation des différents éléments génomiques. La **conjugaison** bactérienne est un des processus de dissémination de gènes entre cellules indépendamment de la reproduction [Lawley et al., 2004]. Elle implique le transfert d'ADN de cellule à cellule par contact avec établissement du canal de conjugaison, veritable "pont" moléculaire traversant les membranes cellulaires. Ce phénomène initialement décrit comme une spécificité plasmidique, a ensuite été caractérisé chez les ICE [Wozniak and Waldor, 2010].

Les systèmes de conjugaison sont constitués de trois sous-systèmes essentiels [Lawley et al., 2004; Smillie et al., 2010] :

- Le transférosome (système de sécrétion de type IV : T4SS) qui est responsable de la déstabilisation de l'enveloppe cellulaire et de la synthèse du pilus.
- Le relaxosome, complexe protéique qui agit sur l'ADN au niveau de l'origine de transfert oriT. Les enzymes du relaxosome sont responsables de l'ouverture du plasmide au site oriT et du désenroulement de l'ADN grâce à des hélicases.
- Les protéines de couplage (T4CP) qui s'accrochent au transférosome et intéragissent avec la relaxase.

La conjugaison suit les étapes suivantes [Lawley et al., 2004] : i) Synthèse d'un pilus par la cellule donneuse et son intéraction avec la membrane de la cellule receveuse, suivie de la création d'un canal entre les deux membranes (*mating pair formation*). ii) Intervention du relaxosome (et en particulier de la relaxase), qui coupe l'ADN de l'élément conjugatif au site *nick*. Un des deux brins est alors "déroulé" et passe dans la cellule receveuse. iii) La réplication des deux molécules d'ADN simple brin dans les cellules par les machineries moléculaires des hôtes.

La réplication des éléments conjugatifs est homologue à celle des plasmides de type RC. On distingue les éléments conjugatifs codant leur propre système T4SS, des éléments mobilisables possédant oriT, relaxases et éventuellement un système T4CP [Smillie et al., 2010]. Ces derniers peuvent être mobilisés et transférés à des cellules receveuses via les systèmes T4SS et T4CP d'autres éléments conjugatifs.

Il est remarquable que des modules fonctionnels d'ADN liés à la conjugaison et reliés d'un point de vue évolutif se retrouvent sur des plasmides de bactéries Gram-positif et Gram-négatif, ainsi que sur des transposons de type ICE et IME [Francia et al., 2004]. De même, l'homogénéité relative des séquences oriT identifiés sur certains plasmides et sur les ICE chez les bactéries Gram-positif et Gram-négatif suggère une origine ancienne [Francia et al., 2004]. Ces homologies suggèrent que ces éléments génétiques ont évolués à partir d'un groupe de gènes et/ou de motifs d'ADN ancestraux qui ont divergé en s'adaptant à des situations particulières [Francia et al., 2004].

1.8 Fluidité du matériel génétique

"Rather than being a well-designed blueprint, a genome appears to be a temporary community of potentially mobile genes that essentially act selfishly" [Mochizuki et al., 2006].

Les différentes structures génomiques bactériennes et la relative diversité des mécanismes moléculaires impliqués dans les réplication, ségrégation et maintenance du génome laissent entrevoir la plasticité du génome bactérien. Les génomes sont soumis à des processus de modification et de transfert d'ADN internes, faisant intervenir un ou plusieurs éléments génétiques et réplicons, ou externes, impliquant plusieurs génomes bactériens (ou un génome et de l'ADN extra-cellulaire). Les échanges inter-génomiques sont généralement rassemblés sous la désignation de Transfert Horizontal de Gènes (THG) et regroupent les processus de conjugaison (introduit dans la partie précédente), de transduction [Miller, 2004], de transformation [Miller and Day, 2004], ou encore les Gene Transfer Agents plus récemment décrits. Les THG tiennent une part très importante dans l'évolution des génomes bactériens [Wiedenbeck and Cohan, 2011] et font intervenir des mécanismes moléculaires liés à la réplication, ségrégation et maintenance du génome. Ces mécanismes tiennent donc une place centrale dans la définition et l'évolution des génomes, notamment pour ce qui est de l'intégration et de la stabilisation dans l'organisme bactérien des différents éléments génomiques. Les mécanismes de recombinaison, en particulier, interviennent largement dans les échanges et inversions entre éléments génomiques. Les homologies structurelles et fonctionnelles partagées par les différentes classes de recombinases montrent des similitudes et indiquent de probables origines communes à ces molécules. Il semble de plus que l'on retrouve ces similitudes de manière plus générale (par exemple, avec les systèmes Par ou les systèmes conjugatifs). Ainsi se dessine une continuité entre les différents réplicons et éléments génétiques

retrouvés chez les bactéries ainsi que parmi les mécanismes moléculaires à l'œuvre dans leur maintenance.

Chapitre 2

Les réplicons extra-chromosomiques essentiels

2.1 Découverte et premières caractérisations de génomes multipartites bactériens

Le modèle du génome bactérien que nous avons introduit dans le chapitre précédent repose essentiellement sur l'architecture du génome d'*E. coli*, formé d'un unique chromosome circulaire et, éventuellement, de plasmides. Cette organisation proposée dans les années 1950 [Wollman et al., 1956] et mise en évidence par autoradiographie du génome de *E. coli* [Cairns, 1963], fut plus tard retrouvée lors de l'étude de génomes additionnels, *i.e.*, *Bacillus subtilis*, et s'imposa alors comme la règle chez les bactéries. Ce modèle n'est cependant pas valide pour toutes les bactéries. Les génomes bactériens peuvent comporter plusieurs réplicons de type chromosomique (**génome multipartite**), et les réplicons bactériens, chromosomes et plasmides, pouvent être circulaires ou linéaires [Baril et al., 1989; Casjens, 1998; Kolstø, 1997].

Les premières preuves de génomes multipartites bactériens apparurent avec la caractérisation par électrophorèse en champ pulsé du génome de *Rhodobacter sphaeroides* à la fin des années 1980, qui permit de réaliser la seconde carte physique de génome bactérien, après celle d'*E. coli* [Suwanto and Kaplan, 1989a,b]. Ces résultats mirent en évidence que deux des réplicons de *R. sphaeroides* portent des gènes codant des ARNr et sont strictement essentiels pour la bactérie. Ces réplicons furent alors désignés par les termes de chromosome primaire (pour le plus grand) et chromosome secondaire [Suwanto and Kaplan, 1989b]. Ces approches méthodologiques ont par la suite mis en évidence la présence de multiples réplicons essentiels chez *Leptospira interrogans* [Zuerner et al., 1993], *Brucella* [Michaux et al., 1993], *Pseudoalteromonas haloplanktis* [Lanoil et al., 1996], *Burkholderia* [Rodley et al., 1995] et Vibrio cholerae [Trucksis et al., 1998].

Depuis, les progrès des techniques de séquençage ont permis d'obtenir les séquences complètes de milliers de génomes bactériens révélant ainsi une grande diversité quant au nombre de chromosomes et de plasmides par complément génomique, à leurs tailles, et à leurs topologies, ainsi qu'une hétérogénéité de répartition de ces caractéristiques parmi les différentes lignées bactériennes [Casjens, 1998; Mackenzie et al., 2004]. L'étude du catalogue, toujours plus important, des génomes dont la séquence complète est disponible, ainsi que la compréhension du développement bactérien *in vivo* et *in vitro* montrent qu'il existe non seulement des réplicons accessoires de taille et de complexité génomiques proches de celles des chromosomes mais que la limite entre réplicon essentiel et réplicon accessoire peut être difficile à définir [Mackenzie et al., 2004].

Pour certains auteurs, le statut des réplicons dépend avant tout des caractéristiques structurales de ceux-ci [Harrison, 2011]. Des études des structures d'*ori* de chromosomes secondaires ont montré de grandes similarités avec les *ori* de mégaplasmides, rapprochant ces deux types de réplicons d'un point de vue évolutif et fonctionnel. Différents modules génétiques ou gènes spécifiques, typiques des plasmides (tels que les gènes impliqués dans la conjugaison), sont retrouvés sur les chromosomes secondaires. Une autre ambiguïté provient de la distinction classique, conforme au dogme du chromosome bactérien unique, entre les différents réplicons essentiels d'un génome bactérien multipartite : on parle alors d'un chromosome *primaire* et de chromosome secondaire ou mégaplasmide [MacLellan et al., 2004] pour les **R**éplicons **Extra-Chromosomiques Essentiels (RECE)** sur la base de caractéristiques telles que la taille des réplicons, le nombre de gènes ou la présence de gènes essentiels.

2.2 Diversité actuelle des génomes multipartites bactériens

Les occurrences de génomes multipartites dans les différentes lignées bactériennes sont hautement variables. Les génomes multipartites sont disséminés de façon irrégulière à travers la phylogénie des bactéries (Figure 2.1). Les Protéobactéries semblent être leur phylum de prédilection car on les observe beaucoup plus fréquemment chez les Alpha-, Bêta- et Gamma-protéobactéries. Chez les Alpha-protéobactéries, jusqu'à 10 genres bactériens différents comportent au moins une espèce à génome multipartite. Par contre, le phylum des Firmicutes n'en comporte que peu (deux génomes multipartites répertoriés) au regard du nombre d'espèces dont le génome a été complètement séquencé (745 espèces dans 90 genres taxonomiques). Néanmoins, la présence de réplicons secondaires essentiels a été relevée dans la plupart des lignées bactériennes majeures telles que les Spirochètes (Leptospira), les Bacteroidetes (Prevotella), les Firmicutes (Butyrivibrio et Clostridium), ou encore les Cyanobactéries (Anabaena et Cyanothece) (Table 2.1; Annexe A). Chez les Protéobactéries, 19 genres sur 231 contiennent au moins une espèce possédant un génome multipartite, pour seulement deux des 90 genres de firmicutes, soit quatre fois moins. Cela traduit peut-être simplement un échantillonnage déséquilibré vers les organismes pathogènes ou symbiotiques pour des raisons médicales et/ou économiques. L'écologie et le génome d'espèces telles que Candidatus Chloroacidobacterium thermophilum [Garcia Costas et al., 2012] ou Ilyobacter polytropus [Sikorski et al., 2010] en sont des contre-exemples, témoignant de l'effort récent pour étendre l'acquisition de données de séquences aux bactéries de l'environnement et à l'ensemble des bactéries, et plus généralement du vivant.



FIGURE 2.1: Distribution des génomes multipartites (bleu) dans le domaine Bacteria.

L'existence de RECE n'est pas corrélée avec la présence ou l'absence de plasmide au sein d'un génome. Ils peuvent être :

- observés ponctuellement au sein d'une lignée (cas des Alteromonadales où seules deux espèces de *Pseudoalteromonas* possèdent un génome multipartite,
- retrouvés chez un grand nombre d'espèces d'une même lignée sans être ubiquiste (cas des Rhizobiales où de nombreuses espèces (*Brucella, Ochrobactrum, Rhizobium* et *Sinorhizobium*), mais pas toutes (*Bradyrhizobium, Mesorhizobium...*) ont des génomes multipartites,
- ou encore présents dans toutes les espèces d'une lignée (cas des Vibrionales).

La taille des RECE varie considérablement, le plus petit (0,27 Mb) étant trouvé chez Leptospira biflexa et le plus grand (3,7 Mb) chez Burkholderia gladioli. Enfin, même si la majorité des RECE a une topologie circulaire, certains, dans les génomes de Agrobacterium et de Cyanothece, sont linéaires.

Les espèces à génomes multipartites possèdent des écologies variées, pouvant être pathogènes (*Brucella, Vibrio, Burkholderia*), des symbiotes de plantes (*Burkholderia, Rhizobium*) ou d'animaux (*Vibrio, Photobacterium*), avoir une écologie marine (*Pseudoalteromonas, Vibrio, Anabaena*) ou être extrêmophiles (*Deinococcus*), par exemple.

2.3 Essentialité

L'essentialité d'un réplicon dépend directement de celle des gènes qu'il code. Ainsi, selon la définition d'Ochman [Ochman, 2002], un réplicon est considéré comme un chromosome s'il porte au moins un gène considéré comme essentiel à la vie de la bactérie, présent en copie unique dans le génome. La notion de gène essentiel pour un organisme relève du concept de "génome minimal" où 100 à 300 gènes sont absolument indispensables à la survie et au développement de l'organisme bactérien [Glass et al., 2006; Koonin, 2000. Ces gènes codent majoritairement des protéines impliquées dans la réplication de l'ADN, la transcription des gènes, la traduction, et l'utilisation du carbone et de l'énergie [MacLellan et al., 2004]. Le statut de chromosome s'applique ainsi au second réplicon de Rhodobacter sphaeroides car il porte des gènes d'ARNr et de certaines enzymes du métabolisme primaire Suwanto and Kaplan, 1989b]. Le statut de chromosome a également été accordé aux réplicons de Butyrivibrio proteoclasticus [Kelly et al., 2010], Nocardiopsis dassonvillei [Sun et al., 2010] et *Ilyobacter polytropus* [Sikorski et al., 2010], car ils portent des gènes codant des ARNr et des ARN de transfert (ARNt). Le chromosome secondaire de Rhizobium radiobacter (anciennement Agrobacterium tumefaciens) possède la plupart des gènes codant des régulateurs transcriptionnels et des transporteurs ABC en plus de gènes d'ARNt et ARNr [MacLellan et al., 2004]. De même, le second réplicon essentiel de Sinorhizobium meliloti porte une part significative des gènes impliqués dans le métabolisme cellulaire et l'exploitation des sources de carbone et d'azote présentes au niveau du sol [MacLellan et al., 2004].

La distinction entre chromosome et plasmide n'est cependant pas aussi tranchée et ne peut reposer sur la seule présence (chromosome) ou absence (plasmide) de gènes "marqueurs" sans copie sur un autre réplicon du genome. Certains réplicons plasmidiques portent des gènes essentiels [Yeoman et al., 2011], les plus fréquemment observés étant des opérons ARNr [Baril et al., 1992; Moran et al., 2004; Pohlmann et al., 2006; Salanoubat et al., 2002]. À l'inverse, il existe des RECE qui n'en codent pas [Heidelberg et al., 2000; Nascimento et al., 2004]. Par ailleurs, de nombreux réplicons bien que plasmidiques contribuent significativement à la *fitness* de leur hôte [Krone et al., 2007; López-Guerrero et al., 2012; MacLellan et al., 2004], le terme *fitness* désignant ici ce qui a rapport avec les capacités de survie et de développement d'un organisme bactérien à court, moyen et long terme, dans son (ou ses) habitat(s) naturel(s). Du fait de la difficile tâche de mesurer la contribution d'un gène ou d'un réplicon à la *fitness* d'un organisme [Mackenzie et al., 2004], il devient délicat de différencier chromosomes et plasmides uniquement sur la contribution d'un réplicon à la *fitness* de l'organisme bactérien.

En règle générale, la distribution en gènes essentiels est largement déséquilibrée en faveur d'un réplicon principal (**chromosome primaire** ou **chr**) [Mackenzie et al., 2004], même si certains RECE font exception [Choudhary et al., 2004, 2007; Egan, 2005]. Parmi les RECE recensés, on observe différents cas de figure (Tableau 2.1), empêchant l'inférence directe de règles quant à la distribution de gènes essentiels en nombre et type sur les RECE.

TABLE 2.1: Caractéristiques essentielles des RECE selon le genre de leur hôte.

Source : génomes complets disponibles dans RefSeq au 23/11/2013. M : nombre de génomes multipartites rapporté au nombre total de génomes séquencés pour un genre bactérien donné. T et R : présence de gènes codant des ARNt et ARNr. Références : références bibliographiques ou numéros d'accession des BioProjects (NCBI); la première référence correspond à la première caractérisation du génome multipartite du genre quand elle est disponible.

Lignée	\mathbf{M}	\mathbf{T}	\mathbf{R}	Autres caractéristiques	Références

Alphaproté	obactér	ries			
A grobacterium	4/4	Oui	Oui	Présence de <i>minCDE</i> , de gènes liés à la réplication de l'ADN, au métabolisme, à l'infection et de régulateurs transcriptionnels	[Allardet-Servent et al., 1993; Goodner et al., 2001; Wood et al., 2001], [MacLel- lan et al., 2004; Slater et al., 2009]
A stic cacaulis	1/1	Oui	Oui	-	PRJNA55641
Brucella	18/18	Oui	Oui	Présence de l'opéron <i>minCDE</i> et différents gènes d'ARNt-synthétases	[DelVecchio et al., 2002; Jumas-Bilak et al., 1998; Paul- sen et al., 2002], [MacLellan et al., 2004
Ochrobactrum	1/1	Oui	Oui	Présence d'homologues de $minC$ et $minD$	[Chain et al., 2011; Slater et al., 2009]
Paracoccus	1/1	Oui	Non	Présence de nombreux gènes d'ARNt- synthétases ou méthyltransférases	PRJNA58187
Rhodobacter	4/5	Oui	Oui	Présence de nombreux gènes typiquement chromosomiques liés au métabolisme; Sous-représentation des gènes liés à la di- vision cellulaire	[Kiley and Kaplan, 1988; Mackenzie et al., 2001; Suwanto and Kaplan, 1989b], [Choudhary et al., 2004, 2007; Mackenzie et al., 2007, 2004]
Sinorhizobium	1/1	Non	Non	Présence de <i>minCDE</i> , de gènes liés à l'infection et au métabolisme, et de régulateurs transcriptionnels; RECE si- milaire aux mégaplasmides des Rhizobia- ceae	[Kaneko et al., 2000], [MacLellan et al., 2004; Slater et al., 2009]
Sphingobium	2/2	Oui	Oui	Présence de gènes codant diverses en- zymes essentielles, des sous-unités d'une ARN polymérase et des gènes impliqués dans le métabolisme et l'adaptation à l'en- vironnement	[Copley et al., 2012; Nagata et al., 2010]

Bêtaprotéobactéries

Burkholderia	39/40	Oui	Oui	Présence de nombreux gènes impliqués dans le métabolisme primaire et la synthèse d'acides aminés; Répartition des gènes essentiels variable selon les espèces. Pour l <i>B. cenocepacia</i> , ratio (chromosome/RECE-1/RECE-2) en gènes essentiels de l'ordre de 250/50/50	[Komatsu et al., 2003; Nierman et al., 2004; Songsivilai and Dharakul, 2000], [Guo et al., 2010]
--------------	-------	-----	-----	--	--

Lignée	\mathbf{M}	\mathbf{T}	\mathbf{R}	Autres caractéristiques	Références
Cupriavidus/ Ralstonia	6/13	Oui	Oui	Présence d'un gène d'une sous-unité de la polymérase et d'un facteur d'élongation avec des paralogues sur le chromosome. Présence de nombreux gènes impliqués dans l'infection des plantes et la forma- tion du flagelle.	[Salanoubat et al., 2002], [Amadou et al., 2008; MacLellan et al., 2004; Pohlmann et al., 2006]
Variovorax	2/3	Non	Oui	chromosome largement favorisé en gènes essentiels	[Han et al., 2011]

Gammaprotéobactéries

Pseudo- alteromonas	2/3	Non	Non	présence de gènes essentiels $hisE$ et $gcpE$, ainsi que des gènes du métabolisme de l'histidine sur le RECE de P . halo- planktis; génomes multipartites des deux espèces très similaires	[Médigue et al., 2005], [Qin et al., 2011]
Vibrio	28/28	Oui	Oui	Nombreux gènes impliqués dans la réparation de l'ADN, gènes impliqués dans le métabolisme énergétique, gènes de protéines ribosomales, d-sérine désaminase et thréonyl-tRNA synthétase ; quasi majorité des gènes d'ARNr portée par le chromosome	[Trucksis et al., 1998; Yamaichi et al., 1999], [Heidelberg et al., 2000; MacLellan et al., 2004; Tagomori et al., 2002]
Aliivibrio	1/1	Oui	Oui	Génome proche du celui des Vibrio	[Hjerde et al., 2008]
Photobacterium	1/1	Oui	Oui	RECE en moyenne 25% plus grand que ceux de <i>Vibrio</i>	[Vezzi et al., 2005]

Acidobactéries

<i>Candidatus</i> Chloroacido-	1/1	Non	Oui	Présence de gènes essentiels impliqués dans la chlorophototrophie	[Garcia Costas et al., 2012]
bacterium					
thermophilum					

Actinobactéries

No cardiops is	1/1	Oui	Oui	-	[Sun et al., 2010]

Bacteroidetes

Prevotella	3/6	Oui	Oui	_	PRJNA47507,
					PRJNA51377,
					PRJNA65091,
					PRJNA163151

Chloroflexi

Sphaerobacter	1/1	Oui	Oui	-	[Pati et al., 2010]

Cyanobactéries

Cydhosdotorros								
Anabaena	1/3 Non Oui Présence de génes d'ARNt synthétases é		[Wang et al., 2012]					
				ARNr méthyltransférases				

Lignée	\mathbf{M}	\mathbf{T}	\mathbf{R}	Autres caractéristiques	Références
Cyanothece	1/6	Non	Non	Présence de génes d'ARNt synthétases; Nombreux gènes singuliers du RECE linéaire	[Welsh et al., 2008]
Thermobaculum	1/1	Non	Non	-	[Kiss et al., 2010]

Deinococcus-Thermus

Deinococcus	1/7 Oui		Oui	Présence de gènes d'ARNt synthétases; Nombreux gènes spécialisés dans l'adap-	[White et al., 1999], [MacLellan et al., 2004]
				tation à l'environnement	

Firmicutes

Butyrivibiro	1/2	Oui	Oui	Présence des gènes essentiels $acpD$ et $msrA$; gènes essentiels présents sur un plasmide de <i>B. proteoclasticus</i>	[Kelly et al., 2010], [Yeoman et al., 2011]
Clostridium	1/62	Oui	Non	-	[He et al., 2010]

Fusobactéries

Ilyobacter	1/1	Oui	Oui	-	[Sikorski et al., 2010]
------------	-----	-----	-----	---	-------------------------

Spirochètes

Leptospira	7/7	Non	Non	Présence déshydrogéna: synthétase) mialdéhyde de	de se), et ehyd	n gltE asd ro-géna	dh 3 ((aspa use)	(NAI (glutama artate	DH ate se-	[Ren et al., 2003; Zuer- ner et al., 1993], [Bu- lach et al., 2006; Nasci- mento et al., 2004; Pi- cardeau et al., 2008]
------------	-----	-----	-----	---	--------------------------	-----------------------------	----------------------------	----------------------------	------------------	--

Outre les RECE annotés, le statut de certains réplicons est ambigu. Bien que classés parmi les (méga)plasmides sur la base de caractéristiques structurales ou fonctionnelles, ils semblent jouer un rôle de RECE ou paraissent en posséder certaines caractéristiques. Ainsi, chez les Alphaprotéobactéries, différents plasmides de *Rhizobium elti* (Rhizobiales) présentent de nombreuses caractéristiques (stabilité, essentialité...) de RECE [Landeta et al., 2011], ainsi que chez *Azospirillum* (Rhodospirillales), bactéries promouvant la croissance de diverses plantes [Acosta-Cruz et al., 2012; Wisniewski-Dyé et al., 2011]. Différents plasmides des *Roseobacter* seraient également des RECE sur la base de caractéristiques structurales (*ori*, pourcentage en G+C) des réplicons [Petersen et al., 2013].

2.4 Régulation et intégration des RECE dans le cycle cellulaire

La présence d'un chromosome additionnel dans un génome ajoute des contraintes à l'organisme et requiert des mécanismes moléculaires supplémentaires permettant son intégration dans le cycle cellulaire [Venkova-Canova, 2011]. Par opposition à un plasmide, un RECE doit être répliqué une unique fois en un moment précis du cycle cellulaire et doit être activement ségrégé lors de la division cellulaire [Egan, 2005].

Le modèle le plus populaire d'intégration d'un RECE dans le génome est l'adaptation des systèmes de réplication et ségrégation d'un mégaplasmide au cycle cellulaire [Heidelberg et al., 2000; MacLellan et al., 2004]. Les mécanismes les mieux décrits sont ceux de V. cholerae où l'hypothèse plasmidique d'origine du RECE est postulée [Egan and Waldor, 2003]. Il a été montré que le chromosome et le RECE de V. cholerae se répliquent une fois par cycle cellulaire [Srivastava and Chattoraj, 2007; Stokke et al., 2011]. La réplication débute après celle du chromosome et termine en synchronisation avec lui Rasmussen et al., 2007; Stokke et al., 2011], et implique des régulateurs communs [Demarre and Chattoraj, 2010; Egan, 2005]. Le RECE de V. cholerae possède deux gènes spécifiques, rtcB codant RtcB, protéine initiatrice de la réplication du RECE des Vibrionaceae et Photobacteriaceae à l'instar de DnaA pour la réplication du chromosome, et rtcA intervenant dans la régulation de RctB [Duigou et al., 2006]. RtcB est de plus régulée par le chromosome par fixation à des motifs spécifiques localisés sur ce dernier Baek and Chattoraj, 2014], ce qui démontre l'implication du chromosome dans la réplication du RECE et, partant, l'intégration du RECE dans le cycle cellulaire et le génome stable de V. cholerae. La réplication du RECE est inhibée par une séquence adjacente à l'ori, inc, et semble aussi sous le contrôle de DnaA et Dam de par la présence de boîtes DnaA et de motifs GATC à l'origine [Egan et al., 2006; Saint-Dic et al., 2008]. Interviennent également des mécanismes de handcuffing, typiques des plasmides [Egan, 2005; Egan and Waldor, 2003; Zakrzewska-Czerwińska et al., 2007]. Le RECE possède son propre système de partition ParA2/ParB2, essentiel à la maintenance du RECE, qui interagit (via ParB2) avec rtcA et RtcB, liant ainsi réplication et ségrégation du RECE [Yamaichi et al., 2011]. Les mécanismes moléculaires de résolution de dimère de RECE font intervenir les mêmes résolvases XerC et XerD que pour le chromosome en le liant au cycle cellulaire par l'intermédiaire de l'action de FtsK [Val et al., 2008]. FtsK est, pour le RECE comme pour le chromosome de V. cholerae guidé par des motifs KOPS Val et al., 2008].

Le modèle de réplication/ségrégation de V. cholerae met en évidence des adaptations génomiques spécifiques du RECE : présence de régulateurs et protéines supplémentaires spécifiques, motifs structuraux de régulation, qui semblent être le fruit de la modification d'un mégaplasmide chez l'ancêtre commun des Vibrio [Heidelberg et al., 2000]. Cependant, ce modèle est actuellement le seul décrivant en quelques détails les spécificités génomiques permettant l'intégration des RECE dans le cycle cellulaire, et ne peut de facto être directement généralisé à tous les RECE.

Les autres mécanismes d'intégration d'un réplicon additionnel dans le cycle cellulaire concernent principalement les spécificités de la partition chez les génomes multipartites des Burkholderiales, avec des adaptations des systèmes parABS des RECE en comparaison de ceux des chromosomes et des plasmides [Dubarry et al., 2006; Livny et al., 2007; Passot et al., 2012]. En particulier, ces adaptations impliquent une séquence *parS* spécifique des RECE et une régulation des systèmes de partition des RECE de *B. ceno-cepacia* par opposition à un système de partition plasmidique classique [Dubarry et al., 2006].

2.5 Origine évolutive

Quelles sont donc les structures génétiques ancestrales expliquant la diversité actuelle des réplicons secondaires essentiels ? D'un point de vue théorique, plusieurs hypothèses peuvent être envisagées :

Deux hypothèses sont principalement favorisées :

- H1 La scission chromosomique. Un unique réplicon se divise en plusieurs réplicons essentiels, faisant potentiellement intervenir des recombinaisons site-spécifiques (cf. §1.7). La coupure expérimentale d'un chromosome en deux molécules qui sont ensuite transmises à la descendance a été décrite [Guo et al., 2003; Itaya and Tanaka, 1997].
- H2 L'adaptation d'un réplicon accessoire avec enrichissement progressif en gènes essentiels chromosomiques [Mackenzie et al., 2004; Moreno, 1998]. Cette hypothèse implique l'action de divers processus de recombinaisons génétiques entre chromosomes et plasmides, dont certains ont été identifiés et caractérisés [Guo et al., 2010; Maida et al., 2014; Slater et al., 2009]. Cette hypothèse est actuellement la plus généralement retenue pour rendre compte de l'existence de la plupart des génomes multipartites, en particulier de ceux des Protéobactéries [Bavishi et al., 2010].

D'autres mécanismes ont été proposés :

- h1 Capture d'un chromosome externe [Mackenzie et al., 2004; Moreno, 1998]. Il est alors nécessaire d'imaginer la formation d'une forme intermédiaire hybride, contenant plus d'un complément génomique.
- h2 Une réplication chromosomique inégale et/ou imparfaite faisant apparaître deux réplicons différents [Moreno, 1998], ce qui devrait se refléter dans la présence des régions dupliquées relativement abondantes dans le génome.
- h
3 Des mutations indépendantes sur des chromosomes d'espèces polyploïdes [Moreno, 1998], men
ant à une différentiation des chromosomes .

Même si les hypothèses h1, h2 et h3 sont vraisemblables, il n'existe pas à notre connaissance d'indice génomique indiscutable en leur faveur. Les hypothèses les plus susceptibles d'expliquer la formation des génomes multipartites les plus étudiés, avec les éléments les soutenant, sont synthétisés Table 2.2.
TABLE 2.2: Origine évolutive des RECE d'après la littérature.

 ${\rm H}$: hypothèse(s) évolutive(s) des RECE.

Lignée	н	Origine de réplication	Autres caractéristiques

Alphaprotéobactéries

A stic cacaulis	?	Pas d'étude publiée de la structure génomique d'Asticcacaulis			
Agrobacterium	H2	Origine de réplication de type $repABC$ [Slater et al., 2009]	Signatures génomiques proches de celles des plasmides; ParA de type plasmidique [MacLellan et al., 2004; Slater et al., 2009].		
Brucella	H2/ (H1/ h2/ h3)	Pour certains auteurs, l'ori du RECE de B. melilentis est similaire à celle du chromo- some [DelVecchio et al., 2002], ce qui se- rait compatible avec les hypothèses H1,h2 et h3. Un locus repABC est aussi présent [Ma- cLellan et al., 2004]; l'ori du RECE semble donc être homologue à celles de différents mégaplasmides et RECE de Rhizobiaceae [Paulsen et al., 2002].	Les ParA des RECE sont des homologues de protéines plasmidiques [MacLellan et al., 2004] et de nombreux indices témoignent aussi d'un passé plasmidique proche de ce- lui des autres mégaplasmides de la famille [Slater et al., 2009].		
Ochrobactrum	H2	Origine de réplication de type <i>repABC</i> [Chain et al., 2011].	Les génomes d' <i>Ochrobactrum</i> et de <i>Bru-</i> <i>cella</i> sont très proches phylogéniquement et génétiquement.		
Paracoccus	?	-	Fréquents transferts entre les chromosomes et plasmides dans le genre <i>Paracoccus</i> [Maj et al., 2013]. Certains plasmides des Rhodobacterales (dont fait partie <i>Paracoc- cus</i>) possèdent des plasmides <i>DnaA-like</i> ca- ractéristiques [Petersen et al., 2011]		
Rhodobacter	H1/ H2/ h2	Présence d'un opéron <i>repABC</i> [Mackenzie et al., 2007]	Large duplication ancestrale entre le chro- mosome et le RECE [Choudhary et al., 2004]. Expression génique faible [Macken- zie et al., 2007].		
Sin or hizobium	H2	Origine de réplication de type <i>repABC</i> [Barnett et al., 2001; Slater et al., 2009]	RECE très proches des mégaplasmides des autres Rhizobiaceae [Slater et al., 2009]		
Sphingobium	H2	Origine de réplication à proximité des gènes $parA$ et $parB$ ainsi que de $repA$ [Copley et al., 2012]	Biais de distribution important des gènes essentiels en faveur du chromosome [Copley et al., 2012].		

Bêtaprotéobactéries

Burkholderia	H2	Origine de réplication possédant des ca- ractéristiques plasmidiques : proximité de <i>parA</i> et <i>parB</i> . Cependant : proximité de <i>dnaG</i> , <i>polA</i> et <i>rpoD</i> (transcription) absent du chromosome [Holden et al., 2004]	Nombreux réarrangements et phénomènes de translocation intragénomique [Guo et al., 2010; Komatsu et al., 2003]. Peu de gènes essentiels près des <i>ori</i> des RECE [Na- gata et al., 2005]. Systèmes de partition <i>pa-</i> rABS [Dubarry et al., 2006; Morrow and Cooper, 2012; Passot et al., 2012]
Cupriavidus, Ralstonia	H2	ori typique de mégaplasmide : proximité de $repA$ dont le produit est similaire à des protéines plasmidiques [MacLellan et al., 2004; Salanoubat et al., 2002]	De même que pour <i>Burkholderia</i> , protéines parA homologues de protéines plasmi- diques [MacLellan et al., 2004; Passot et al., 2012].
Variovorax	(H2)	L'origine de réplication n'a pas été identifiée dans le RECE [Han et al., 2011]	Distribution en gènes essentiels suggestive d'une origine de type H2 [Han et al., 2011].

Gammaprotéobactéries

Pseudo-	H2	Origine de réplication proche de celle du plas-	La	réplication	serait	unidirectionnelle
alteromonas		mide R1 [Médigue et al., 2005].	[Mé	digue et al., 2	2005].	

Lignée	\mathbf{H}	Origine de réplication	Autres caractéristiques	
Lignée Aliivibrio, Photobacterium, Vibrio	H H2	Origine de réplication ori de type plasmidique malgré des ca- ractéristiques uniques, présence de boîtes DnaA et motifs IHF, et mécanismes de régulation spécifiques chez V. cholerae [Ma- cLellan et al., 2004]. Pas d'étude décrivant spécifiquement les ori des Photobacterium et	Autres caractéristiques Protéine ParA proche des protéines plasmi- diques [Thompson et al., 2004]. Présence d'intégron typique des plasmides chez V. cholerae [Heidelberg et al., 2000], mais pas chez les autres Vibrionaceae (intégron sur le chromosome). Des transferts entre le	
		Aliivibrio	chromosome et le RECE ont été mis en évidence. Présence importante de transpo- sases et transposons sur les RECE de <i>P.</i> <i>profundum</i> et de <i>A. salmonicida</i> [Chen et al., 2003; Egan and Waldor, 2003; Kir- kup et al., 2010; MacLellan et al., 2004]	

Acidobactéries

Candidatus	?	-	Distribution en gènes identique à celle du
Chloroacido-			chromosome [Garcia Costas et al., 2012].
bacterium			RECE plus sensible aux réarrangements
thormophilum			génétiques que le chromosome [Garcia Co-
mermophinum			stas et al., 2012].

Actinobactéries

No cardiops is	?	-	Le génome de N. alba (non présent dans
			la base de données) a été séquencé et ne
			comporte qu'un unique chromosome [Qiao
			et al., 2012]

Bacteroidetes

Prevotella	?	_	Très grande diversité génomique au sein du
1 100000000	•		genre <i>Prevotella</i> [Purushe et al., 2010].

Chloroflexi

Sphaerobacter	?	-	-
Thermobaculum	?	_	-

Cyanobactéries

Cyanothece	?	Les <i>ori</i> n'ont pu être déterminées sur le chro- mosome ou le RECE par les méthodes de biais en GC [Welsh et al., 2008]	RECE linéaire
Anabaena	?	-	Présence de gènes liés aux réplicases sur un plasmide [Kaneko et al., 2001]; Traces de prophages insérés [Wang et al., 2012]. Exis- tence d'espèces à génome monopartite avec des plasmides similaires en taille au RECE [Kaneko et al., 2001].

Deinococcus-

Thermus

Deinococcus	H2/ (h3)	ori à proximité de $parAB$; gène rep très simi- laire à un gène plasmidique [MacLellan et al., 2004]	D. radiodurans est typiquement polyploïde [White et al., 1999]. Des ressemblances existent entre le BECE de D. radiodurans
		2004j	et un mégaplasmide de <i>Thermus thermo-</i> philus [Omelchenko et al., 2005]

phages présents chez les Spirochètes [Picar-

deau et al., 2008].

Lignée	Η	Origine de réplication	Autres caractéristiques
Firmicutes			
Butyrivibiro	H2	ori de type plasmidique [Yeoman et al., 2011]. Un plasmide de <i>B. proteoclasticus</i> possède de nombreuses protéines de la réplication de type chromosomique [Yeoman et al., 2011]	très petit (0,3 Mb) [Kelly et al., 2010]
Clostridium	?	-	-
Fusobactéri	\mathbf{es}		
Ilyobacter	?	-	-
Spirochètes			
Leptospira	H2	$ \begin{array}{ c c c } Présence de boîtes DnaA à proximité de l'ori \\ [Ren et al., 2003]. Opéron de partition parAB \\ et gène rep à proximité d'ori [Picardeau et al., \end{array} $	Les caractéristiques génomiques de <i>Leptos-</i> <i>pira</i> suggèrent que le RECE (ainsi que cer- tains plasmides) ont évolué à partir de pro-

Concernant l'origine des différents génomes multipartites identifiés, on peut se demander si leur apparition est antérieure à celle de la lignée bactérienne (espèce, genre, famille les possédant ou si, au contraire, ces phénomènes sont récents et postérieur à l'émergence de la lignée (spéciation).

2008].

La dispersion des génomes multipartites dans les différentes lignées bactériennes et à des niveaux variables de différentiation (espèce, genre, famille) suggère que **plusieurs** apparitions indépendantes de génomes multipartites ont eu lieu au cours de l'évolution des génomes bactériens. Ceci implique que l'émergence de génome multipartite est relativement "facile" et s'effectue par l'action de mécanismes génétiques assez classiques. Il a été proposé que tous les RECE de Rhizobiaceae descendraient d'un réplicon extrachromosomique de type mégaplasmide, à réplication RepABC qui a capturé des gènes du chromosome, et dont des marqueurs géniques différents (minCDE, repABC, par exemple) sont retrouvés dans chacune des lignées de Rhizobiaceae [Slater et al., 2009]. Alors que le mégaplasmide s'est intégré stablement dans le chromosome chez certaines espèces (Bradyrhizobium), il a conservé son état de mégaplasmide chez d'autres ou bien a été tout simplement éliminé (Bartonella) [Slater et al., 2009]. Chez les Vibrionales, les RECE semblent avoir coexisté avec le chromosome avant la diversification en différentes lignées, dont Vibrio, Aliivibrio, Photobacterium et espèces de cette famille [Thompson et al., 2004]. De même chez les Burkholderiales, l'analyse des systèmes parABS [Passot et al., 2012] met en évidence des similarités entre RECE et mégaplasmides des Burkholderia et Ralstonia, ce qui suggère un ancêtre commun plasmidique aux RECE de cette famille de bactéries. B. rhizixinica fait cependant figure d'exception, son génome ne comportant qu'un seul chromosome et deux plasmides [Lackner et al., 2011]. Néanmoins, B. rhizixinica en tant que pathogène endosymbiotique, possède un génome réduit et on peut faire l'hypothèse que cette espèce a opéré une réduction de son matériel génomique par un mécanisme similaire à celui décrit par Moreno [Moreno, 1998] (cf. paragraphe suivant). À l'inverse, les RECE présents chez Deinococcus, Clostridium, Prevotella, Anabaena, Cyanothece et Rhodobacter semblent

être d'apparition relativement récente car des espèces du même genre ou de la même famille possédent un génome monopartite.

2.6 Rôle

À ce jour, considérant la diversité d'organisation des génomes multipartites, il est difficile d'établir un rôle ou une fonction précise à cette architecture génomique. Il semble cependant improbable que l'organisation multi-chromosome du génome n'apporte aucun bénéfice à l'hôte, surtout si on postule que l'intégration et la stabilisation d'un second chromosome dans un génome représente un coût évolutif. Différentes tendances semblent se dessiner quant aux possibilités d'amélioration de la *fitness* apportées à l'hôte par un génome multipartite.

- ► Les RECE peuvent être des réservoirs de gènes accessoires, spécifiques d'un certain type d'écologie. Chez les protéobactéries, les fonctions codées par les RECE sont majoritairement en lien avec des voies métaboliques spécifiques ou des mécanismes d'infection ou d'interaction avec différents hôtes [Galardini et al., 2013].
- ► Les RECE peuvent permettre une réplication plus rapide du génome ce qui est un avantage pour des espèces à taux de croissance élevé comme V. cholerae en conditions favorables [Yamaichi et al., 1999]. Une réplication rapide peut aussi résulter par le lancement d'un cycle de réplication avant la fin du cycle cellulaire précédent [Stokke et al., 2011] comme chez Escherichia coli [Skarstad et al., 1986].
- ► Une structure du génome en plusieurs répliquons stables entraîne, lors de la réplication, une augmentation de l'expression génique par le phénomène de "gene dosage" pour les gènes proches d'ori car dupliqués en premier [Jha et al., 2012], à la fois au niveau du chromosome mais également du RECE. Cela permet de fournir rapidement et en grande quantité les protéines nécessaires au développement de la colonie, et pourrait apporter une nouvelle modalité de régulation de l'expression génique.
- Chez Vibrio cholerae, il a été suggéré que les deux chromosomes, dans certaines conditions, pourraient présenter une différence du nombre de copies, avec pour effet de moduler le niveau d'expression de certains gènes [Heidelberg et al., 2000]. Un cas extrême serait la perte du chromosome ou du RECE qui conduirait à la formation de "cellules drones" favorisant la survie de la population par un taux de sécrétion d'enzymes élevé [Jha et al., 2012].
- ► La perte du RECE a également été proposée comme mécanisme d'adaptation en réponse à des conditions environnementales défavorables dans le cas de Vibrio cholerae [Heidelberg et al., 2000]. Cependant, V. cholerae n'est pas capable de survivre à plus d'un cycle cellulaire dans un tel cas [Yamaichi et al., 2007]. De plus, les chromosomes et RECE de V. cholerae N16961 sont asservis l'un à l'autre par des systèmes Toxine-Antitoxine, entraînant la dégradation du chromosome en cas de perte du RECE [Yuan et al., 2011].

- ▶ Les génomes multipartites peuvent être une solution structurelle aux génomes de grande taille, long et difficile à se répliquer en un seul réplicon. Cependant l'existence de chromosomes bactériens de plus de 10 Mb prouve que la taille d'un génome ne peut seule expliquer la formation d'un RECE.
- Avoir un génome multipartite peut être un avantage par l'augmentation de la surface d'échange du génome avec le cytoplasme, ce qui permet, à taux d'échange constant, de réduire la taille des cellules et ainsi d'avoir une structure cellulaire plus adaptée à des environnements pauvres en nutriments [Morita, 1988; Stouthamer and Kooijman, 1993].
- ► L'acquisition de gènes essentiels par un plasmide peut être vue comme un mécanisme (de type PSK) dans le sens où l'organisme "solidifie sa relation" avec un plasmide qui apporte une contribution significative à l'augmentation de sa *fitness* [Slater et al., 2009].
- ► Les RECE semblent dans de nombreux cas être plus plastiques que les chromosomes. Ils présentent un taux global d'expression génique plus faible et/ou portent relativement moins de gènes et sont donc moins contraints sur le plan évolutif [Choudhary et al., 2004; White et al., 1999]. Chez *R. sphaeroides*, par exemple, les séquences des RECE sont plus divergentes que celles des chromosomes, ce qui semble indiquer une différence des vitesses d'évolution des deux réplicons [Choudhary et al., 2007]. Cette différence peut s'expliquer par le fait que le RECE possède une expression génique plus faible que le chromosome, ainsi que des séquences intergéniques longues et serait donc plus enclin à subir des réarrangements génomiques fréquents. Les RECE peuvent ainsi servir d'"atelier évolutif" en étant des réservoirs des gènes à évolution rapide et moins exprimés, ou exprimés sporadiquement, ce qui est favorable aux bactéries dans certains milieux [Bavishi et al., 2010; Cooper et al., 2010; Galardini et al., 2013; Morrow and Cooper, 2012].
- ▶ De plus, en facilitant les recombinaisons *inter* et *intra*-chromosomiques, la structure multipartite peut participer à l'augmentation de la diversité des séquences [Mackenzie et al., 2004], à l'instar du phénomène décrit chez les Fungi [Croll and Mcdonald, 2012].
- ► Enfin, chez les organismes endosymbiotes, un génome multipartite peut refléter un état de transition vers la réduction génomique dans lequel l'hôte se "déleste" d'une partie des gènes essentiels en les transférant à ses plasmides, qui sont éliminés par la suite [Moreno, 1998]. Brucella a été proposé comme exemple de cette configuration [Wattam et al., 2009].

La présence d'un état multipartite, conservé à travers le temps au cours de la diversification de différents groupes bactériens, reflète le succès évolutif de cette architecture génomique et son importance pour l'adaptation et l'évolution des organismes bactériens. Cela pourrait traduire des modalités d'évolution génomique spécifiques à certains milieux écologiques [Slater et al., 2009], comme par exemple dans les cas des Rhizobiaceae et des Burkholderiales où l'on observe de nombreux génomes multipartites et qui renferment une proportion importante d'espèces symbiotiques ou parasites des plantes. Cette architecture génomique pourrait aussi constituter une alternative efficace (ou neutre) à une structure classique monopartite du génome. Le cas de *R. sphaeroides* illustre la dernière hypothèse : *R. capsulatus* du même genre, possède un génome monopartite sans que cela semble être un avantage [Choudhary et al., 2007; Tichi and Tabita, 2001].

2.7 Critères d'identification des RECE

Les propriétés caractéristiques des RECE sont :

- l'Essentialité Élément les distinguant fondamentalement des plasmides.
- l'Intégration dans le cycle cellulaire impliquant une coordination avec le chromosome et une synchronisation avec la division cellulaire
- Contribution à la *fitness* de la bactérie d'une manière ou d'une autre la *fitness*, sur le court ou long terme.

Classiquement les RECE ont été majoritairement identifiés par leur caractère essentiel, notamment par la présence de gènes codant des ARNr et ARNt. Cependant la présence/absence de tels gènes est loin d'être suffisante pour classer un réplicon parmi les RECE.

2.7.1 Modèle du "chromid"

Récemment, l'appellation "chromid" a été introduite par Harrisson et al. [Harrison et al., 2010] pour décrire les RECE à partir du contenu en G+C et de la stratégie d'utilisation des codons synonymes (*Relative Synonymous Codon Usage*; RSCU), le type d'origine de réplication, et les protéines ParA/B [Harrison, 2011], et a depuis été reprise par plusieurs auteurs lors d'études ponctuelles de génomes de protéobactéries [Acosta-Cruz et al., 2012; Galardini et al., 2013; Maj et al., 2013; Petersen et al., 2013; Ramírez-Bahena et al., 2014; Wisniewski-Dyé et al., 2011].

Selon ce modèle :

- Les *chromid* possèdent des systèmes de réplication et de maintenance proches de ceux des plasmides.
- Les chromid ont des compositions en G+C et RSCU proches de ceux des chromosomes (et relativement plus que des plasmides) ce qui suggère qu'ils ont co-habité avec le chromosome pendant relativement plus longtemps que ne le font des plasmides classiques [Harrison et al., 2010, Fig. 1].
- Les chromid portent des gènes essentiels, mais aucun des 284 gènes identifiés du génome-coeur n'est présent sur les chromid de Burkholderia étudiés [Harrison, 2011]. Ils portent de plus de nombreux gènes accessoires (relativement plus que les chromosomes), qui sont conservés au sein d'un genre bactérien.

Les gènes communs à l'ensemble des membres de la famille bactérienne (Burkholdériales) étudiée par Harrison *et al.* sont relativement conservés alors qu'ils ne le sont que faiblement quand ne sont considérés qu'un genre (*Burkholderia*) ou qu'une espèce (*B. cenocepacia*), indiquant une plasticité des RECE accrue par rapport aux chromosomes dans les génomes étudiés [Harrison et al., 2010, Fig. 2]. Enfin, l'ordre des gènes sur les chromid est très peu conservé en comparaison des chromosomes à des niveaux taxonomiques supérieurs au genre [Harrison, 2011]. Ces constatations rejoignent les résultats des travaux de Bavishi *et al.* [Bavishi et al., 2010] où les longueurs relatives des alignements significatifs des RECE sont comparées à celles des chromosomes pour les génomes de souches de la même espèce ou d'espèces différentes, ce qui suggère une différence de vitesse d'évolution entre chromosomes et RECE [Bavishi et al., 2010].

2.7.2 Autres modèles

Plusieurs études ont tenté de classer les différents types de plasmides et réplicons accessoires. Des critères structurels comme la présence de complexe de mobilisation/conjugaison ainsi que les caractéristiques des différentes relaxases peuvent servir pour discriminer les réplicons accessoires [Garcillán-Barcia et al., 2009]. Il est cependant à souligner que les RECE, par leur rôle positif probable sur la *fitness* de la bactérie, n'ont pas besoin de système de maintenance "égoïste" de type plasmidique [Smillie et al., 2010]. D'autres travaux proposent de classer les plasmides selon les gènes des modules de réplication [Jensen et al., 2010; Petersen, 2011]. Ces approches peuvent dans un premier temps sembler pertinentes pour la caractérisation des RECE, ceux-ci adaptant potentiellement ces protéines afin de s'intégrer dans le cycle cellulaire. Cependant, ces études se focalisent sur un gène ou opéron (repABC) unique et ne prennent pas en compte les mécanismes intervenant dans la ségrégation et la maintenance des RECE.

2.8 Données génomiques de notre étude

La question se posant naturellement à propos de la structure spécifique des réplicons secondaires essentiels est : **Existe-il des gènes, ou distributions de gènes, com-muns à l'ensemble de ces réplicons ?** Différents éléments de réponse, abordés dans ce chapitre montrent qu'il n'existe pas de règle de fixation d'un gène, ou groupe de gènes, particulier sur les réplicons secondaires accessoires permettant d'expliquer leur transformation en réplicon essentiel. Les analyses des chapitres suivants ont donc pour objectif de caractériser d'éventuels biais de répartition de gènes des STIG à partir de l'analyse de l'intégralité des séquences complètes des génomes bactériens multipartites référencés dans **RefSeq** [Pruitt et al., 2007] à la date du 23/11/2013 (*cf.* Annexe A). Notre jeu de données comprend 2016 génomes bactériens et 1267 plasmides isolés, soit 2016 chromosomes, 2781 plasmides et 129 RECE selon les annotations de RefSeq. Les distributions de la taille et du nombre de gènes (les deux étant hautement corrélés) présents sur les réplicons bactériens (Figure 2.2) met en évidence de nettes différences entre RECE, plasmides et chromosomes primaires.



(A) Tailles des génomes



(B) Nombre de gènes

FIGURE 2.2: Distributions des tailles (2.2a) et nombres (2.2b) de gènes des réplicons.

Les boites délimitent les premiers et troisièmes quartiles (Q) des distributions, les limites des "moustaches" inférieures et supérieures représentant respectivement Q1-1.5(Q3-Q1) et Q3+1.5(Q3-Q1). Les points en dehors des moustaches sont représentés par des croix. **1** : chromosomes des génomes multipartites. **2** : totalité des chromosomes. **3** : RECE. **4** : plasmides. **5** : différences des tailles (2.2a) ou nombres de gènes (2.2b) entre chromosome et RECE au sein d'un génome. **6** : différences des tailles (2.2a) ou nombres de gènes (2.2b) entre chromosome et plasmide au sein d'un génome.

Ces facteurs ne sont pas suffisants pour caractériser les différents réplicons : il existe des plasmides aussi grands que des RECE, et des RECE aussi grands que des chromosomes. Enfin, la répartition très inégale des génomes multipartites dans les différents groupes et genres bactériens (Figure 2.3), pour des raisons évolutives ou historiques, crée un biais dans les données, biais dont il faudra tenir compte dans l'analyse.



(A) Génomes multipartites à travers le domaine bactérien. Total :133



(B) Genres bactériens ayant au moins un génome multipartite. Total :**31**



FIGURE 2.3: Répartition par lignée des génomes bactériens disponibles au 23/11/2013 dans la base de données RefSeq classés selon la taxonomie bactérienne en vigueur (http://www.bacterio.net/).

Seuls les génomes complets ont été pris en compte. Les génomes sont organisés selon le phylum de leur hôte sauf pour les Protéobactéries pour lesquels la classe a été utilisée. La catégorie "Autres" regroupe les phylum faiblement représentés dans RefSeq : Aquificae, Chlorobi, Deferribacteres, Fibrobacteres, Fusobactéries, Genmatimonadetes, Nitrospirae, Planctomycètes, Thermodesulfobactéries, Thermotogae et Verrucomicrobia.

2.9 Nature, origine et fonctionnement des RECE : une problématique ouverte

Jusqu'à maintenant, la problématique de ce que sont les RECE a été abordée principalement au cours d'études ponctuelles de génomes individuels et parfois, d'espèces bactériennes. Plusieurs tentatives de définition ont été faites [Mackenzie et al., 2004], la plus complète et récente étant celle de Harrison [Harrison, 2011], sans pour autant clôturer le débat sur la nature chromosomique vs. plasmidique des RECE. En effet, l'étude de Harrison ne s'attache qu'à la comparaison des RECE aux chromosome et plasmide(s) présents dans le même génome, et est majoritairement focalisée pour ce qui est de l'étude approfondie, sur une seule espèce (*Burkholderia cenocepacia*) d'un seul genre (*Burkholderia*) d'une seul famille (Burkholderiaceae). Ces résultats ne peuvent donc pas être directement généralisés à l'ensemble des RECE bactériens. Les travaux de Harrison et al. [Harrison et al., 2010; Harrison, 2011] apportent néanmoins de précieuses observations, dont l'interprétation va au-delà de l'introduction de l'appellation de "chromid". En effet, la définition du chromid revient à dire que les RECE fonctionnent comme des chromosomes mais ne peuvent être appelés chromosome parce qu'ils sont plasmidiques à l'origine. La question devient alors : **Qu'est-ce qu'un chromosome**?

Les différents éléments présentés dans ces deux premiers chapitres mettent en évidence une continuité génomique entre plasmides, RECE et chromosomes, et révélent l'existence de mécanismes moléculaires et structures génomiques communs à ces différents types de réplicons. Cette similitude est d'autant plus affirmée qu'il devient difficile de discerner si certains réplicons sont des (méga)plasmides ou des chromosomes. Ainsi la définition même de RECE n'est pas appropriée, la distinction entre réplicons bactériens pouvant au final se résumer à deux éléments : leur essentialité pour l'organisme les hébergeant et leurs intégration et stabilisation dans le cycle cellulaire, leur rôle en étant une conséquence. Du point de vue de l'évolution du matériel génétique, les bactéries à génome multipartite représentent de fait une collection de données génomiques cruciale pour appréhender les nature et importance des forces évolutives mises en œuvre dans l'organisation du matériel génétique et la complexification et l'adaptation des génomes bactériens. Or, les études faites jusqu'à présent ne permettent pas de caractériser clairement et de manière générale ce que sont un plasmide, un chromosome ou un RECE parmi les réplicons bactériens.

Les travaux de cette thèse portent sur la caractérisation des génomes multipartites par une **étude globale sans** *a priori* **de l'ensemble des génomes bactériens et données disponibles dans les bases de données publiques** dont l'objectif est d'identifier des tendances générales des mécanismes génétiques impliqués dans l'apparition et la stabilisation des architectures génomiques multipartites.

Chapitre 3

Stratégie d'étude

Les Systèmes de Transmission de l'Information Génétique (STIG) sont les mécanismes génétiques permettant la réplication, la ségrégation et la maintenance des réplicons au cours des générations. Le passage d'un génome monopartite à multipartite (ou inversement) doit logiquement passer par une adaptation des STIG du génome afin de permettre l'intégration de la réplication et de la ségrégation d'un chromosome additionnel (ou sa disparition) dans le cycle cellulaire et vis-à-vis de chromosomes préexistants. Il apparaît alors pertinent de caractériser les réplicons et génomes bactériens selon les STIG qu'ils utilisent. Dans cette étude, nous recherchons plus spécifiquement à mettre en contraste les RECE par rapport aux chromosomes primaires (et chromosomes uniques) et aux "vrais" plasmides afin, premièrement, de discriminer efficacement les différentes catégories de réplicons et secondairement, d'identifier les systèmes-clés de l'émergence et de la stabilisation des génomes multipartites.

3.1 Les protéines des STIG : variables explicatives des réplicons

Les chapitres 1 et 2 ont présenté les STIG des réplicons bactériens, ainsi que la conception actuelle de l'architecture du génome bactérien. Il semble alors que les types de réplicon, et les espèces et écologies des organismes les hébergeant sont fonctionnellement liés à leurs STIG. Selon notre postulat d'adaptation des STIG dans les génomes multipartites, des différences de distribution des gènes liés aux STIG entre génome mono- et multi-partites doivent être détectables par une analyse bioinformatique. Nos données de départ étant l'ensemble des génomes bactériens séquencés disponibles dans les bases de données publiques, nous annotons dans un premier temps l'ensemble des gènes liés aux STIG dans ces génomes. En prenant ces gènes comme attributs des réplicons, nous analysons ensuite ce jeu de données pour identifier et caractériser d'éventuels biais de distribution entre les réplicons, afin, à partir de ces résultats, de distinguer les différents types de réplicons. Enfin, nous relions ces biais à des réalités biologiques et génomiques. Deux types d'attributs sont utilisés pour caractériser les réplicons bactériens :

- les groupes d'homologues des protéines liées au STIG,
- les différentes fonctions de ces protéines.

Dans le premier cas, la mesure rapprochant deux protéines étant l'homologie de séquence, la proximité de deux réplicons traduira un lien évolutif potentiel. Dans le second cas, les réplicons seront identifiés selon leur fonctionnalité.

3.2 Fondement de la génomique comparative

Trois modalités d'analyse bioinformatique peuvent être distinguées :

- L'analyse de séquences biologiques (protéines, ARN, ADN), "analyse" sous-entendant la comparaison (recherche d'homologie) des séquences.
- La modélisation moléculaire qui consiste à inférer le comportement et/ou l'organisation de structures moléculaires.
- La biologie des systèmes, cherchant à construire des modèles reliant différents niveaux d'information biologique (structural, moléculaire, écologique...).

Un des paradigme fondamental est que l'homologie de séquence est liée à une origine évolutive commune des séquences comparées [Wray and Abouheif, 1998] et, à un degré moindre, à une homologie fonctionnelle [Chikina and Troyanskaya, 2011]. Ces hypothèses permettent, par exemple, d'étudier les relations phylogéniques entre individus ou d'annoter une séquence biologique inconnue par rapport à des séquences biologiques déjà annotées.

L'homologie de deux séquences peut refléter différents mécanismes évolutifs :

• Orthologie [Fitch, 1970] Deux séquences chez deux organismes distincts sont orthologues si elles descendent d'une unique séquence ancestrale présente dans le dernier ancêtre commun aux deux espèces. L'homologie de deux gènes séparés par un événement de spéciation laisse supposer que les gènes ont conservé la fonction du gène ancestral et ont, *a priori*, la même fonction.



(A) **Orthologie** entre les gènes α des espèces A et B.

• Paralogie [Fitch, 1970] Deux g sont paralogues si elles résultent d'une duplication génique. On distingue de plus les In-paralogues si la duplication a eu lieu avant l'événement de spéciation, et les Out-paralogues si la duplication a eu lieu après la spéciation.



(C) **Out-paralogie** entre le gène α de A et le gène β de B.

La duplication génique est souvent suivie d'une divergence fonctionnelle des gènes dupliqués. Un des gènes conserve la fonction du gène ancestral et l'autre diverge pour acquérir une nouvelle fonction [Fitch, 1970, 2000].

• Xénologie Deux gènes sont xénologues si leur homologie traduit une histoire évolutive commune mais, pour l'une des espèces les hébergeant, n'est pas liée au gène présent dans l'ancêtre direct dont ces dernières sont issues. Dans les génomes bactériens, la présence de séquences xénologues témoigne fréquemment d'un transfert horizontal de gènes (THG) inter-réplicons et/ou inter-individus. L'évolution des génomes bactériens étant soumise à un flux important de THG ([Bapteste et al., 2009; Doolittle and Bapteste, 2007]), la détection d'orthologues au sein des différents réplicons est souvent délicate et bruitée par les xénologues.

3.3 Concept du génome-cœur

De nombreuses études mettent en avant la présence d'une collection de gènes homologues chez l'ensemble des membres d'une espèce (par exemple, pour *Salmonella enterica* [Lan et al., 2009] ou *Helicobacter pylori*). Cela a conduit à la proposition de l'hypothèse du **génome-cœur** ou *Core Genome Hypothesis* [Lan and Reeves, 1996]. Cette théorie distingue la fraction du génome (le cœur) partagée par tous les membres de l'espèce, de la fraction auxiliaire qui est seulement présente dans une sous-population. Le génomecœur définit ainsi les spécificités d'une espèce bactérienne. Les gènes du génome-cœur codent principalement les fonctions métaboliques essentielles et les processus cellulaires de transmission d'information [Riley and Lizotte-Waniewski, 2009]. Ainsi, une part importante des gènes des STIG sont inclus dans le génome-coeur des bactéries. Les génes de la partie auxiliaire sont principalement les gènes codant des fonctions métaboliques supplémentaires, utiles dans le contexte de certaines écologies, et sont plus fréquemment associés à des plasmides ou des éléments mobiles [Riley and Lizotte-Waniewski, 2009].

3.4 Fouille de données en génomique

3.4.1 Méthodes analytiques

Avec l'explosion de la production de données génomiques, et en particulier l'avènement de nouvelles technologies de séquençage de l'ADN et des méthodes *chip-seq*, les chercheurs sont confrontés à un afflux massif de données génomiques, transcriptomiques, protéomiques ou physiologiques. Cette avalanche de données est devenue un défit pour les bio-analystes et requiert le développement constant de méthodes analytiques permettant de traiter de façon pertinente (et de stocker) des quantités toujours plus importantes d'information génétique bruitée dans des temps raisonnables.

Les données à analyser sont souvent multivariées, *i.e.*, décrites par une collection de paramètres (taux d'expression génique, présence/absence de gènes, écologie des organismes ...) généralement considérés comme des variables aléatoires. Parmi les méthodes d'analyse multivariée, on distingue les méthodes de statistiques classiques, impliquant des tests d'hypothèses, des méthodes dites de fouille de données (data mining) généralement employées pour analyser des jeux de données massifs et multi-dimensionnels [Izenman, 2008]. Le terme data-mining fait référence à un concept un peu "fourre-tout" englobant des approches de transformation et de préparation des données, de visualisation, d'extraction de connaissance et d'évaluation. Parmi ces méthodes d'analyses, nous pouvons distinguer les **méthodes descriptives**, qui ont pour objectif d'explorer les données (recherche d'observations aberrantes (*outliers*), de structures ou de tendances, sélection de variables...), des **méthodes prédictives** où l'objectif est de construire un modèle à partir des données. Les méthodes de data-mining ont souvent recourt à des concepts statistiques, mais une part importante de ces méthodes englobe les approches dites d'apprentissage (*Machine Learning*). À partir d'un jeu de données, les algorithmes prédictifs de machine learning cherchent à apprendre, à reconnaître des tendances complexes et sont capables de prendre des "décisions" en fonction des données [Han et al., 2012]. Une façon de conceptualiser les problèmes d'apprentissage est d'imaginer qu'est recherché à travers un espace de concepts descriptifs, un ensemble de règles décrivant au mieux un jeu de données [Witten et al., 2011]. Ainsi, un algorithme de machine learning est défini par les différents concepts descriptifs utilisés pour représenter les données, l'ordre dans lequel est parcouru l'espace des concepts et la façon dont sont traités les problèmes de sur-apprentissage (ou *overfitting*) des jeux de données d'entraînement (*training set*) [Witten et al., 2011]. Enfin, nous pouvons distinguer les méthodes d'apprentissage supervisé (classification ou régression), où un ensemble de variables d'entrée (*input*) dont l'état de sortie (*output*) est connu sert à entraîner un

algorithme d'apprentissage afin de traiter des données dont l'état de sortie n'est pas connu, des méthodes d'**apprentissage non-supervisé** (clustering par exemple) où aucune information *a priori* n'est disponible sur l'état de sortie des données.

L'efficacité d'un algorithme peut être mesurée selon différents critères. Dans l'analyse de données biologiques, on peut souligner l'importance de :

- La complexité temporelle et spatiale, c'est-à-dire la faisabilité temporelle et matérielle d'un algorithme. Cela ce traduit concrètement par le temps d'exécution d'un algorithme t, et la taille de la mémoire nécessaire m. Formellement les complexités temporelle et spatiale peuvent être représentées par des fonctions $O_t(Y)$ et $O_m(Y)$ indiquant que t et m vont être proportionnels à Y, où Y s'exprime généralement en fonction du nombre de données à analyser n, et de la dimension des données d (Y = f(n, d)). Les problèmes génomiques impliquant généralement des n et d importants, la complexité temporelle exponentielle $(O(e^{n.d}))$ d'un algorithme le rend souvent inutilisable en pratique. Il est alors important de déterminer si un problème informatique est divisible en plusieurs sous-problèmes autorisant sa parallélisation sur plusieurs processeurs.
- La performance. Les méthodes doivent fournir des résultats pertinents des points de vue statistique et biologique. Différents critères d'évaluation peuvent être utilisés, tels que, par exemple, l'erreur de prédiction ou de généralisation [Izenman, 2008] ou la stabilité ou l'insensibilité à l'ordre des entrées [Andreopoulos et al., 2009]. Une propriété importante des méthodes prédictives est leur capacité à généraliser le modèle construit à partir du training set à des jeux de données inconnus. La robustesse des algorithmes est la capacité à générer des résultats pertinents quand les entrées sont des données "bruitées" dans les cas, par exemple, des valeurs manquantes ou singulières [Andreopoulos et al., 2009; Han et al., 2012, Chap. 8].
- Le nombre de paramètres à définir. Les paramètres que doit définir l'utilisateur d'un algorithme peuvent affecter les résultats d'une manière significative [Andreopoulos et al., 2009]. Afin de minimiser les erreurs dues aux choix de l'utilisateur, il peut être préférable qu'un algorithme utilise un nombre limité de paramètres à définir.
- L'inclusion de variables multicatégories. Il est souvent utile en génomique comparative de pouvoir prendre en compte des variables de différents types : par exemple, des groupes de gènes d'une part, et des critères écologiques ou morphologiques d'autre part.

3.4.2 Notations

On note $E = \{e_1, ..., e_n\}$, l'ensemble E des n éléments e_i avec 1 < i < n de taille |E| = n. On peut alors écrire : $E = \{e_1, ..., e_{|E|}\}$ lorsque n n'est pas connu.

On note $v = (x_1, ..., x_n)$, le vecteur v de taille |v| = n avec x_i , les différentes valeurs de v avec 1 < i < n et x_i appartenant à \mathbb{R} ou \mathbb{N} . v appartient alors à \mathbb{R}^n ou \mathbb{N}^n .

On note de plus $v[i] = x_i$. Pour un ensemble $V = \{v_1, ..., v_{|V|}\}$ de vecteurs v_i de même taille n, on note M^V sa matrice de dimension (|V|, n) telle que $M_{i,j}^V = v_i[j]$, avec 1 < i < |V| et 1 < j < n.

Pour un ensemble de données, le terme **observation** désigne une donnée tirée de cet ensemble. Les observations étant souvent représentées par des vecteurs, le terme **attribut** désigne alors les variables descriptives de ces vecteurs. Pour un ensemble d'observations $E = \{e_1, ..., e_{|E|}\}$ définies par *n* attributs $X = \{X_1, ..., X_n\}$, l'ensemble $V^E = \{v^{e_1}, ..., v^{e_{|E|}}\}$ désigne l'ensemble des vecteurs d'observations où $v^{e_i}[j]$ est égal à la valeur du *j*-ième attribut pour l'observation e_i .

3.4.3 Distances

Pour comparer deux observations, plusieurs distances [Gan et al., 2007] peuvent être calculées selon leurs attributs et les propriétés souhaitées :

Si les observations sont deux ensembles E_1 et E_2 , la distance de Jaccard peut être utilisée :

$$d_{Jaccard}(E_1, E_2) = 1 - \frac{|E_1 \cap E_2|}{|E_1 \cup E_2|}$$
(3.1)

Pour deux observations v_1 et v_2 ayant des attributs numériques, et de taille similaire n, la distance la plus couramment utilisée est la distance Euclidienne :

$$d_{Euclid}(v_1, v_2) = \sqrt{\sum_{1 \le i \le n} (v_1[i] - v_2[i])^2}$$
(3.2)

Cette distance n'est pas forcément pertinente pour comparer des observations de très grande dimension dont une partie n'a aucun attribut en commun. On introduit donc une distance Euclidienne modifiée :

$$d_{Euclid*}(v_1, v_2, Seuil) = \begin{cases} d_{Euclid}(v_1, v_2) & \text{si } \prod_{1 \le i \le n} (v_1[i].v_2[i]) \ne 0\\ Seuil & \text{sinon} \end{cases}$$
(3.3)

Une distance est considérée comme *métrique* si elle respecte les propriétés suivantes :

Symétrie : d(a,b) = d(b,a)Séparation : $d(a,b) = 0 \iff a = b$

Inégalité triangulaire : $d(a, c) \leq d(a, b) + d(b, c)$

Les distances de Jaccard et Euclidienne sont métriques. Cette propriété n'est pas toujours pertinente pour des données de dimension élevée et clairsemées (avec beaucoup d'attributs nuls). La distance *cosine* (non métrique car ne respectant pas la dernière propriété) est alors souvent utilisée afin de se focaliser sur les attributs non nuls [Han et al., 2012]:

$$d_{cosine} = 1 - \frac{\sum_{1 \le i \le n} v_1[i] . v_2[i]}{\sqrt{\sum_{1 \le i \le n} v_1[i]^2} . \sqrt{\sum_{1 \le i \le n} v_2[i]^2}}$$
(3.4)

3.4.4 Méthodes de clustering

Soit un ensemble E. On appelle procédure de clustering une procédure f_{clust} qui structure un ensemble d'observations en k sous-ensembles ou **clusters** constitués des observations présentant des similitudes, sans *a priori* sur la structure :

$$f_{clust}(E) = \{E_1, ..., E_k\}$$
 où : $\bigcup_{E_i \in f_{clust}(E)} E_i = E$ (3.5)

De nombreux algorithmes de clustering existent [Gan et al., 2007]. Les inputs de ces algorithmes peuvent être, entre autres, la distance utilisée. Cependant, un des paramètres clé souvent requis est le nombre k de clusters à former.

3.4.5 Méthodes de classification

Une procédure de classification prend en input un training set $E_{training}$, constitué de différents sous-ensembles (ou classes) d'observations : $E_{training} = \{E_1, ..., E_k\}$, et construit à partir de cet ensemble un modèle permettant d'attribuer à une observation o, une des classes de $E_{training}$. Soit E_{test} un ensemble d'observations, et $E_{training}$ un training set tel que $\forall E \in E_{training}, E \subset E_{test}$. Une procédure de classification $f_{classif}^{E_{training}}$ utilisant $E_{training}$ peut alors être alors définie par :

$$f_{classif}^{E_{training}}(o) = i, \ o \in E_{test} \text{ et } 1 \le i \le k$$

$$(3.6)$$

L'objectif sous-jacent est, pour tout $E_i \in E_{training}$ et tout $o_i \in E_i$, de bâtir un classifieur tel que $f_{classif}^{E_{training}}(o_i) = i$, se comportant de manière stable et robuste, et évitant l'overfitting de $E_{training}$. Différents algorithmes de classification sont présentés plus loin. Une description plus complète de ces méthodes est disponible dans [Han et al., 2012; Witten et al., 2011].

3.4.6 Projection

Une procédure de projection consiste à projeter des observations représentées dans un espace de dimension p, dans un nouvel espace de dimension q avec q < p.

Soit f_P^q la procédure de projection P de données provenant d'un espace Q^p de dimension p dans un espace Q'^q de dimension q. Pour une observation $v = (x_1, ..., x_p)$ décrite par p attributs avec $v \in Q^p$ et $x_i \in Q$, on définit f_P^q par :

$$f_P^q(v) = (y_1, ..., y_q) \text{ où } y_i \in Q'$$
 (3.7)

Ici $Q \in \{\mathbb{R}, \mathbb{N}\}$ et Q' est l'ensemble des réels \mathbb{R} . Pour un ensemble d'observations $V = \{v_1, .., v_{|V|}\}$, on définit M_P^V par :

$$M_P^V = f_P^q(V) = \{ f_P^q(v_1), \dots, f_P^q(v_{|V|}) \} \text{ où } v_i \in V$$
(3.8)

Une procédure de projection peut être utilisée dans le but de représenter des données dans un espace interprétable et visuel. Dans ce cas, q est classiquement fixé à 2 ou 3. Ici, les données sont visualisées dans un espace à deux dimensions en fixant q à 2.

3.4.7 Évaluation

L'efficacité d'une analyse bioinformatique peut être appréhendée de deux manières : i) d'un point de vue informatique et statistique, où l'on mesure la performance des résultats selon différents critères, et ii) d'un point de vue biologique, où l'on cherche à voir si les résultats sont significatifs biologiquement et non un artefact découlant, par exemple, d'un biais initial dans les données (sur-représentation d'une classe, données manquantes...).

3.4.7.1 Performance des classifieurs

L'évaluation de la performance des méthodes de classification se fait *via* l'évaluation de différents indices à partir des résultats apportés par les classifieurs et selon des procédures spécifiques. En considérant les résultats d'un classifieur sur un training set, soient TP et FP, les résultats vrais (T pour "true") annotés positivement (P) et négativement (N), et FP et FN, les résultats faux (F pour "false") annotés positivement (P) et négativement (N), respectivement. Nous pouvons définir les mesures suivantes [Hamel, 2011; Han et al., 2012] :

sensibilité =
$$\frac{TP}{TP + FN}$$

précision = $\frac{TP}{TP + FP}$
spécificité = $\frac{TN}{TN + TP}$ (3.9)
 $erreur = \frac{TP + TN}{FP + FN + TP + TN}$
 $F - measure = \frac{2.precision.sensibilite}{precision + sensibilite}$

Ces estimateurs ne témoignent pas de la capacité d'un classifieur à se généraliser à des données indépendantes de celles du training set utilisé pour sa construction, de bons résultats pouvant simplement témoigner d'un *overfitting* du classifieur sur le learning set initial. La capacité à se généraliser d'un classifieur peut être estimée en partitionnant le learning set initial en différents E_{test} et $E_{learning}$ et évaluant les performances du classifieur sur les E_{test} formés. Les procédures de validation croisée **cross-validation** sont des procédures suivant ce principe [Hamel, 2011; Han et al., 2012; Witten et al., 2011]. La méthode **Leave-One-Out** (**LOO**) consiste à effectuer *n* classifications sur les *n* données du training set où, à chaque fois, une donnée (ou point) différente est classée en utilisant les n - 1 autres données comme nouveau training set. La méthode **K-fold** consiste à séparer le training set en *K* partitions et à effectuer *K* classifications en testant à chaque fois une partition par rapport aux K - 1 autres. La performance peut alors être mesurée avec, par exemple, les indices décrits. Une forte sensibilité, spécificité et précision (et donc un F-score important), et une faible erreur témoigneront d'un classifieur performant.

Afin d'estimer si deux classifieurs présentent des performances différentes, des procédures de ré-échantillonnage de type **bootstrap** (ré-échantillonnage avec remise du training set initial) peuvent être effectuées où, pour chaque procédure, une nouvelle cross-validation est réalisée et sert à construire une distribution des scores obtenus pour les différents indices (notamment l'erreur de classification). Les distributions obtenues serviront alors à comparer les modèles (par des modèles statistiques de type test d'hypothèses) afin de choisir le plus performant [Hamel, 2011; Han et al., 2012]. Le bootstrap peut aussi servir de procédure de type cross-validation [Han et al., 2012].

Les modèles de cross-validation LOO, K-fold et de bootstrap sont les principales représentantes d'une famille plus large de méthodes plus ou moins bien adaptées aux différents problèmes de classification [Arlot and Celisse, 2010]. Ces modèles sont cependant eux-mêmes soumis à une sélection de paramètres (K par exemple pour K-fold), pouvant être source d'erreur. Une façon d'améliorer les performances d'un classifieur est d'utiliser les méthodes dites d'*Ensemble* dont le principe se réfère à des procédures de ré-échantillonnage des données et/ou des variables, et en combinant les scores obtenus pour chaque échantillon [Han et al., 2012]. Les procédures de *Bagging, Boosting* ou encore les forêts d'arbres aléatoires (*random forest*) font partie des méthodes d'*Ensemble* les plus connues. Le choix d'un modèle particulier étant lui-même source de biais, la sélection d'une procédure d'évaluation de la performance et de la robustesse d'un classifieur est alors un compromis entre faisabilité, complexité et précision [Arlot and Celisse, 2010]. Enfin, selon les classifieurs utilisés, différentes techniques permettent d'évaluer la probabilité d'appartenance d'une donnée à une classe donnée [Rüping, 2004; Wu et al., 2004].

3.4.7.2 Performance des algorithmes de clustering

L'estimation de la performance d'une solution de clustering est généralement soumise à deux types de critères.

Les critères de validation externe [Gan et al., 2007; Han et al., 2012] où les résultats du clustering Cl sont comparés à ceux d'un autre clustering de référence R construit, par exemple, par des experts ou dérivé d'un autre algorithme de clustering

et/ou avec d'autres variables. En général, une mesure de validation externe prendra en compte deux critères :

- l'homogénéité des clusters mesurant la capacité des clusters $C_i \in Cl$ à contenir des données homogènes envers un cluster $R_j \in R$ du clustering de référence.
- l'exhaustivité des clusters mesurant inversement l'homogénéité des $R_j \in R$ envers les $C_i \in Cl$.

Deux critères additionnels peuvent également être pris en compte : la **catégorie bruit** représentant un cluster $C_{bruit} \in Cl$ et incluant les éléments ne correspondant à aucun C_i , et la **préservation des petits clusters**. Différents indices ont été développés afin de comparer deux solutions de clustering [Gan et al., 2007; Rendón et al., 2011]. Pour a, le nombre de paires ayant des points dans le même cluster dans Cl et R, pour b, le nombre de paires où les points sont dans le même cluster dans Cl mais pas dans R et c, inversement, et pour d, le nombre de paires où les points sont dans différents clusters dans Cl et R, on obtient au total M paires où $M = a + b + c + d = \frac{n(n-1)}{2}$. Des indices simples de validation externe peuvent alors être définis [Gan et al., 2007] :

Statistique de Rand :
$$R = \frac{a+d}{M}$$

Coefficient de Jaccard :
$$J = \frac{a}{a+b+c}$$
 (3.10)

Indice de Flokes et Mallows :
$$FM = \sqrt{\frac{a}{a+b} \cdot \frac{a}{a+c}}$$

Les critères de validation interne [Gan et al., 2007; Rendón et al., 2011] mesurant la qualité d'un clustering uniquement à partir des caractéristiques intrinsèques de la partition des données obtenues. En général, ces indices combinent deux concepts [Rendón et al., 2011] : i) la compacité qui mesure la proximité des éléments au sein d'un même cluster (typiquement la variance intra-cluster), et ii) la séparabilité qui mesure la distance séparant deux clusters distincts. De nombreux indices internes ont été développés [Gan et al., 2007; Rendón et al., 2011] et comme illustration, nous pouvons présenter le coefficient silhouette [Rousseeuw, 1987] qui est un estimateur interne classique. Pour un clustering Cl de n données et pour un point i dans un cluster C_j , une mesure de qualité s(i) peut être calculée :

$$s(i) = \frac{(b(i) - a(i))}{Max\{a(i), b(i)\}}$$

avec a(i) et b(i) :

$$a(i) = \frac{\sum_{j \in C_i, j \neq i} d(i, j)}{|C_i - 1|}$$

$$b(i) = \min\{\frac{\sum_{j \in C_k} d(j, i)}{|C_j|} | C_k \in Cl \text{ et } C_k \neq C_i\}$$
(3.11)

où a(i) représente la distance moyenne entre i et tous les autres points de C_j et b(i) est le minimum de la distance entre i et les distances moyennes des points des clusters $C_{k,k\neq i}$. Cette mesure permet d'estimer la pertinence de la classification de chaque point dans un cluster. Elle varie entre -1 et 1. Des valeurs proches de 1 indiquent une pertinence élevée, et inversement, des valeurs négatives indiquent que i est plus proche du centroïde d'un autre cluster. La moyenne des s(i) d'un cluster donne un estimateur de la pertinence du cluster et la moyenne de tous les s(i) procure un estimateur du clustering. Un des problèmes majeurs généralement rencontré dans l'estimation de ces indices est qu'ils ont des complexités temporelles et/ou spatiales difficilement applicables sur des jeux de données très larges. Afin d'estimer la pertinence d'une analyse de clustering, certaines procédures ont été développées pour déterminer si la distribution des données est proche d'une distribution aléatoire, ce qui témoignerait d'une absence de structuration [Gan et al., 2007; Han et al., 2012]. Plusieurs méthodologies et critères de validation relatifs permettent de comparer l'efficacité de différents algorithmes ou paramètres à produire des clusters pertinents d'un point de vue statistique [Gan et al., 2007; Kovács et al., 2005]. Notamment, certains auteurs ont proposé d'étudier la stabilité d'un clustering, l'idée générale étant de réaliser différents clustering d'un jeu de données en utilisant des méthodes de ré-échantillonnage (type bootstrap) et de mesurer la stabilité des clusters obtenus à travers les différents clustering [Fang and Wang, 2012; Hennig, 2007, 2008].

Choix des paramètres Une difficulté généralement rencontrée dans une analyse de clustering est de déterminer le nombre de clusters à considérer, qui n'est pas connu *a priori*. Le choix du nombre optimal de clusters k_{opt} se fait alors en comparant l'efficacité des indices de robustesse pour différents k [Fang and Wang, 2012; Rakhlin and Caponnetto, 2007; Tibshirani and Walther, 2005]. Il est cependant important de ne pas perdre de vue qu'un clustering robuste et performant d'un point de vue statistique peut être complètement dépourvu de sens biologique, celui-ci pouvant correspondre à des artefacts dûs au jeu de données et/ou à l'algorithme utilisé [Hennig, 2007].

3.4.8 Fléau de la dimensionalité

Une des spécificités des données génomiques est qu'elles sont généralement de grande dimension (typiquement nos jeux de données ont des dimensions de l'ordre de 100 à 5000). Or, dans un espace de très grande dimension, le nombre de données disponibles est relativement peu élevé par rapport à la taille de l'espace de recherche. Les données ont alors tendance à se retrouver isolées [Izenman, 2008]. Un des problèmes pratiques qui en découle est que lorsque l'on veut, par exemple, mesurer la distance d séparant quatre points p1, p2, p3 et p4 dans un espace de très grande dimension, en supposant qu'il existe certaines caractéristiques communes entre p1 et p2 dans certaines dimensions, mais aucune entre p1 ou p2 d'une part et p3 ou p4 d'autre part, ni entre p3 et p4, la distance d(p1, p2) ne reflétera pas forcément la proximité de p1 avec p2 en comparaison de d(p3, p4) à cause du bruit potentiellement engendré par les autres dimensions.

Dans le cadre du data-mining, cela se traduit par la nécessité d'avoir d'énormes quantités de données, afin d'explorer au maximum l'espace de représentation des données pour avoir un modèle robuste [Izenman, 2008]. Une façon de contourner ce phénomène est de choisir des modèles et/ou des distances adaptés, comme par exemple un espace de dimension réduit ou des algorithmes spécifiques [Gan et al., 2007].

Cas d'exemple : Soient quatre vecteurs p1 = (5,3,0,0,0,0), p2 = (3,5,0,0,0,0), p3 = (0,0,0,0,0,1), et p4 = (0,0,0,0,1,0). Les distances euclidiennes calculées sont :

p1p2p3*p*4 **p1** 0 p22.830 5.925.92p30 5.925.921.41 0 p_4

On constate que d(p1,2) > d(p3,p4) même si p1 et p2 possèdent des caractéristiques communes (dans les deux premières dimensions) contrairement à p3 et p4 qui n'ont rien en commun.

3.4.9 Graphes

Les graphes sont des outils mathématiques de représentation de données et/ou de concepts qui peuvent être employés sur de nombreux jeux de données génomiques. L'ensemble des similitudes protéine-protéine d'un jeu de séquences est facilement représenté par un graphe (*noeuds* = protéines et *arêtes* = similitudes, qui peuvent être pondérées) qui peut ensuite être utilisé pour la visualisation des données ou faire des analyses de clustering [Brohée and van Helden, 2006; Li et al., 2010]. Les graphes peuvent de plus être utilisés pour représenter des relations entre individus ou génomes comme par exemple, pour représenter un taux de THG ou des relations taxonomiques [Lima-Mendez et al., 2008]. Formellement, un graphe peut être décrit de la façon suivante [Naïm et al., 2011] : Soit $N = \{e_1, \ldots, e_{|E|}\}$, un ensemble non vide d'éléments finis. Un graphe *G* sur *N* peut être défini par la donnée du couple :

$$G = (N, A) \text{ où } A \subset \{(e_i, e_j) | e_i, e_j \in N \text{ et } e_i \neq e_j\}$$

$$(3.12)$$

N est alors nommé l'ensemble des **nœuds** de G. $(e_i, e_j) \in A$ est appelé une **arête** si et seulement si $(e_j, e_i) \in A$ et un **arc**, sinon. Un graphe pondéré, G_w , est défini par $G_w = (N, A, W)$ où W est une matrice carrée de taille |N| * |N|, appelée **matrice d'adjacence**, et W_{ij} , la valeur associée au couple $(e_i, e_j) \in A$ ($W_{ij} = 0$ si $(e_i, e_j) \notin A$). La **distance géodésique** entre deux nœuds est le plus court chemin entre ces nœuds *via* les arêtes (et leurs longueurs) du graphe [Han et al., 2012]. Un graphe peut de plus être décomposé en partitions de noeuds où, à l'intérieur d'une partition, les noeuds sont indépendants et non connectés par des arêtes ou des arcs. De tels graphes sont appelés des **graphes multipartites**. En particulier, pour un ensemble d'observations $E = \{e_1, ..., e_{|E|}\}$ définies par un ensemble d'attributs $X = \{X_1, ..., X_{|X|}\}$ et représentées par un ensemble de vecteurs $V^E = \{v_{e_1}, ..., v_{e_{|E|}}\}$, on peut définir un graphe bipartite pondéré $G_w = (N, A, W)$ de la manière suivante :

$$\begin{cases} N = E \cup X \\ A = \{(e_i, X_j) | v^{e_i}[j] \neq 0\} \\ W_{ij} = v^{e_i}[j] \end{cases}$$
(3.13)

où $e_i \in E$ et $X_j \in X$.

Les graphes possèdent des méthodes d'analyse propres. Les méthodes dites de *détection* de communautés font référence aux méthodes de partitionnement de graphes et sont assimilables aux méthodes de clustering [Coscia et al., 2011]. Une communauté de nœuds peut être vue comme un ensemble de nœuds partageant des caractéristiques communes au sein d'un réseau. Différents critères peuvent alors servir à évaluer le taux de similitude entre différents nœuds [Coscia et al., 2011].

3.4.10 Sources de biais

Tout problème de data-mining peut être vu comme le calcul d'un estimateur $\hat{\theta}$ d'une certaine statistique θ . Les biais de l'estimateur peuvent alors être définis par :

$$Biais(\hat{\theta}) \equiv E[\hat{\theta}] - \theta \tag{3.14}$$

Dans une analyse de machine learning, les sources de biais possibles sont les suivantes [Witten et al., 2011] :

- les différents concepts descriptifs utilisés pour représenter les données,
- l'ordre dans lequel est parcouru l'espace des concepts,
- la façon dont sont traités les problèmes de sur-apprentissage.

La première source de biais est concrètement exprimée par le choix et l'exhaustivité des données génomiques. Dans le cas de notre étude, ces données sont représentées par les réplicons choisis, les protéines sélectionnées en rapport avec les STIG et, dans un second temps, les clusters de protéines obtenus. La deuxième source de biais vient du choix des méthodes utilisées pour les analyses (clustering, classification, régression). Enfin, la troisième source de biais est la sélection des paramètres des méthodes pour éviter qu'ils ne soient pas trop spécifiques aux données. Ainsi, l'interprétation des résultats analytiques devra se faire en toute connaissance des différentes sources potentielles de biais.

3.5 Pipeline analytique de notre étude

La constitution de la *query set database* et de la base de données des réplicons bactériens, la construction de clusters de protéines homologues des STIG et la procédure de cleaning sont décrites dans le Chapitre 4. Les visualisations et clustering des réplicons et des génomes sont décrits dans les Chapitres 5 et 6, et les analyses supervisées dans le Chapitre 7.



FIGURE 3.1: Pipeline analytique.

Chapitre 4

Construction de clusters de protéines homologues des STIG

4.1 Récupération des données brutes : les protéines des STIG des génomes

La collecte des données représente la première étape du pipeline de l'analyse (Figure 4.1). Deux sous-ensembles de données brutes peuvent être distingués : i) les séquences génomiques des réplicons et ii) les protéines des STIG.



FIGURE 4.1: Procédure de récupération des données brutes.

Obtenir une collection quasi-exhaustive des réplicons bactériens séquencés est une tâche aisée car ils sont accessibles dans les bases de données publiques majeures dont celles du NCBI. Par contre, nous avons à faire face à plusieurs difficultés pour la construction d'une base de données des protéines liées aux STIG :

- Sélection d'un ou plusieurs système(s) d'annotation protéique parmi ceux proposés par les bases de données existantes.
- Sélection des annotations spécifiques aux STIG.

 Obtention des données les plus exhaustives possibles pour chaque fonctionnalité des STIG, pour chaque type de réplicon (plasmide et chromosome) et pour toute la diversité taxonomique bactérienne.

4.1.1 Principales sources publiques de protéines annotées

- NCBI (National Center for Biotechnology Information) rassemble des données de séquences d'acides nucléiques et de protéines accessibles via internet. Les bases de données RefSeq [Pruitt et al., 2007] et GenBank [Benson et al., 2008] sont gérées par le NCBI et constituent des collections de séquences (génomes, gènes, protéines...) parmi les plus importantes disponibles. Les protéines sont annotées partiellement selon leur fonction (vérifiées ou hypothétiques) mais aucun formalisme d'annotation n'est proposé. Protein Clusters regroupe différents clusters de protéines homologues annotées par fonction selon les catégories fonctionnelles des COG (clusters of orthologous groups) [Koonin, 2003], et CDD, rassemble des protéines ou alignements de protéines classés selon leurs motifs structurels [Marchler-Bauer et al., 2007].
- KEGG (Kyoto Encyclopedia of Genes and Genomes) est une base de données regroupant des données génomiques, chimiques et systémiques annotées au niveau fonctionnel [Kanehisa et al., 2012]. En particulier, les données des génomes complètement séquencés sont hiérarchisées en fonction de leurs propriétés chimiques et systémiques (KEGG BRITE hierarchy), ce qui, concrètement, produit des groupes de protéines regroupées par annotations fonctionnelles.
- ACLAME (A CLAssification of Mobile genetic Elements) est une base de données dédiée aux données génomiques provenant des MGE (Mobile Genetic Element, terme défini par ACLAME) rassemblant des données de phages, plasmides et transposons [Leplae et al., 2010]. Des clusters de protéines sont accessibles et organisés par familles annotées par des termes de Gene Ontology (GO) [Ashburner et al., 2000] ou PhiGO (système propre d'ontologie).
- **PATRIC** (PathoSystems Resource Integration Center) est un système d'information libre dont le but est de fournir un support à l'analyse des différents pathogènes bactériens [Wattam et al., 2014]. Elle comprend énormément de génomes bactériens de toutes les espèces, et rassemble les annotations de GO fournies par différentes méthodes d'annotations.
- Pfam est une large collection de familles de protéines qui sont représentées par des alignements de séquences multiples et des modèles de Markov, et selon leurs domaines fonctionnels [Finn, Bateman, Clements, Coggill, Eberhardt, Eddy, Heger, Hetherington, Holm, Mistry, Sonnhammer, Tate, Punta, 2014].
- **TIGRFAM** est une base de données de familles de protéines similaire à Pfam [Haft et al., 2003].
- **PDB** (Protein Data Bank) est une base de données de protéines et séquences nucléotidiques comportant des informations structurales (structures 3D) [Rose et al., 2013].

4.1.2 Construction de la base de données requête

Les bases de données KEGG et ACLAME ont été sélectionnées pour la constitution de la base requête de données protéiques. Leur avantage est de proposer des structures d'annotation plus rigoureuses que celles du NCBI. Les annotations proposées par Pfam et TIGRFAM quant à elles sont principalement liées à la fonction des domaines identifiés sur les protéines et permettent difficilement d'identifier de façon exhaustive des groupes de protéines clairement liées aux STIG.

Le système de hiérarchie BRITE a été utilisé dans la base de données KEGG afin d'identifier **68** groupes d'orthologues affiliés aux différentes fonctions protéiques d'intérêt (Chapitre 1). Ce système permet d'organiser différents objets biologiques (notamment des familles de protéines) fonctionnellement (Figure 4.2). Ces 68 groupes d'orthologues (Annexe B) réunissent un total de **43.757** protéines.



FIGURE 4.2: Choix des termes KEGG BRITE pour la constitution de la base requête de données des protéines liées fonctionnellement aux STIG.

91 familles ACLAME ont été selectionnées selon leurs annotations fonctionnelles (GO et PHI) (Table 4.1) et regroupent au total **3.847** protéines (Annexe C). Le choix des familles ACLAME a été effectué de façon semi-automatique. Différentes familles liées aux STIG ont d'abord été sélectionnées manuellement ce qui a permis d'identifier les annotations pertinentes. Les protéines potentiellement liées aux STIG ont alors été identifiées automatiquement parmi les 18.228 familles de protéines plasmidiques, puis confirmées manuellement.

Accession	Description
go 0003677	DNA binding
575	plasmid partitioning protein family ParB/Spo0J
go :0015616	DNA translocase activity
576	plasmid partitioning protein family ParM
go :0000146	microfilament motor activity
go :0007059	chromosome segregation
go :0015616	DNA translocase activity
go :0007059	chromosome segregation
go :0016887	ATPase activity
go: 0030541	plasmid partitioning
go: 0051302	regulation of cell division
phi :0000196	plasmid copy number control

TABLE 4.1: Annotations utilisées pour la selection des Familles ACLAME.

L'ensemble des **47.604** (43.757 + 3.847) protéines de référence de KEGG et ACLAME est noté P_{ref} . L'annotation d'une protéine $p \in P_{ref}$ selon sa famille de référence est désignée par Ann(p). On distingue donc $Ann_{KEGG}(p)$ de $Ann_{ACLAME}(p)$. Les ensembles des familles d'orthologues de KEGG et des familles ACLAME sont notés Cl_{KEGG} et Cl_{ACLAME} , respectivement. Enfin, pour tout $C_i \in Cl_{KEGG}$ et $C_j \in Cl_{ACLAME}$ on désigne par $Ann_{KEGG}(C_i)$ et $Ann_{ACLAME}(C_j)$ l'annotation d'une famille KEGG ou ACLAME donnée.

4.1.3 Récupération des séquences protéiques à partir de RefSeq

L'ensemble des séquences complètes de réplicons disponibles a été récupéré le 23/11/2012via le site FTP de RefSeq. Les séquences protéiques (confirmées ou hypothétiques) codées par les gènes de **5.458** génomes ont été extraites pour un total de **6.903.452** protéines hypothétiques.

4.1.4 Sources de biais possibles

La combinaison des données de KEGG et ACLAME permet de regrouper des protéines annotées d'origines chromosomique et plasmidique. Différentes sources de biais sont possibles :

- Certains groupes fonctionnels peuvent être trop spécifiques à un groupe d'espèces.
 C'est le cas notamment de Hda, DiaA, SeqA, EzrA, SlmA ou RacA qui sont spécifiques de la réplication de *E. coli* et *B. subtilis* notamment, et ne sont pas généralisables à des groupes distants de bactéries.
- À l'inverse, certaines familles protéiques (connues ou non) peuvent ne pas être représentées. C'est le cas par exemple de CtrA, Noc ou de YabA. Le choix de ne

pas sélectionner ces protéines dépend du fait que celles-ci ne sont pas incluses dans les ensembles de protéines liées aux STIG identifiées précédemment. Ces protéines étant trop spécifiques de certains groupes bactériens, il n'a pas été développé de protocole exclusif d'extraction de leurs séquences. Cela risque de ne pas permettre la discrimination ou l'identification des tendances pour certains groupes de réplicons dont trop peu de leurs STIG seraient représentés dans les protéines sélectionnées.

- Il existe un biais d'échantillonnage des familles taxonomiques bactériennes parmi les génomes de réplicons (Figure 2.3). Ce biais peut donner trop de poids aux représentants des familles sur-représentées, et inversement pour les groupes très minoritaires. Il devra en être tenu compte.
- Enfin, d'éventuelles erreurs d'annotation des protéines peuvent fausser les résultats.
 Ce dernier biais, malheureusement, est difficilement estimable dans le cadre de notre étude.

4.1.5 Recherche d'homologues

Parmi les méthodes permettant d'évaluer l'homologie entre séquences protéiques ou nucléotidiques, on peut distinguer celles qui utilisent un alignement des séquences et celles dites "alignment-free sequence analysis" qui répertorient le nombre de mots (ou k-mers) communs entre deux séquences.

Le calcul de distances évolutives entre séquences implique classiquement un alignement des séquences, les distances étant estimées à partir des taux de mutation ponctuelle ou des phénomènes d'insertion/délétion observables à partir de l'alignement. Dans certains cas, des phénomènes extrêmes (inversions, recombinaisons multiples) peuvent biaiser les résultats de ces méthodes, qui ne produiront pas d'alignement cohérent malgré l'existence d'homologie entre les séquences.

Parmi les méthodes d'alignement de séquences (Table 4.2), on peut séparer celles qui effectuent un alignement global cherchant à aligner tous les résidus des séquences-requête, des méthodes d'alignement local où seulement des parties de séquences peuvent être alignées, ces dernières étant plus intéressantes dans le cas des séquences divergentes. On distingue les méthodes présentant des solutions algorithmiques **optimales**, des méthodes dites **heuristiques**, où les résultats sont plus ou moins approchés de l'optimum. On peut de plus distinguer les algorithmes permettant d'aligner des séquences deux à deux, de ceux qui produisent des alignements multiples. Enfin, certains algorithmes requièrent une base de données de séquences à laquelle sont comparées des séquences-requête et produisent des alignements significatifs par rapport à la base de données. TABLE 4.2: Principaux algorithmes d'alignement de séquences de séquences nucléotidiques ou protéiques.

Needleman-(Algorithme de) [Needleman and Wunsch, 1970] Algorithme d'alignementWunschglobal optimal servant à aligner les séquences deux à deux.

Smith-(Algorithme de) [Smith and Waterman, 1981] Algorithme d'alignement localWatermanoptimal servant à aligner les séquences deux à deux.

BLAST (Basic Local Alignment Search Tool) [Altschul et al., 1990] Heuristique de recherche et d'alignement local de séquences, à travers une base de données. Différentes phases composent l'algorithme, les principales étant : la recherche de *k-mer*, la construction d'un alignement local à partir des hits et l'évaluation de l'alignement (statistique de Karlin-Altschul [Korf et al., 2003]). BLAST est une approximation de l'algorithme de Smith et Waterman, et est l'une des méthodes les plus utilisées dans la recherche d'homologie de séquences, le logiciel le plus utilisé étant la suite du NCBI [Camacho et al., 2009]. Le programme PSI-BLAST [Altschul et al., 1997] est une variante de BLAST, qui inclut des procédures itératives de l'algorithme initial afin d'augmenter sa sensibilité à l'identification d'homologues distants.

HMMER [Finn et al., 2011] Suite de logiciels permettant l'analyse de séquences via la création de Modèles de Markov Cachés (HMM pour Hidden Markov Model) [Eddy, 1998]. À partir d'un alignement multiple de protéines de référence, l'algorithme construit un modèle ou profile constitué d'une suite d'états possibles associés à des probabilités de réalisation afin de détecter et d'aligner des séquences homologues potentielles à partir d'une base de données de séquences de référence. hmmsearch, algorithme inclus dans la suite HMMER, présente de meilleures performances (sensibilité et spécificité) que BLAST et PSI-BLAST dans la détection d'homologues [Eddy, 2011]. Jackmmer, un autre algorithme de la suite HMMER est une procédure itérative similaire à PSI-BLAST [Eddy et al., 2013].

MUSCLE [Edgar, 2004] Actuellement un des algorithmes les plus performants et les plus populaires d'alignement multiple de séquences. La première étape de l'algorithme compare les séquences deux à deux par une distance de type *k-mer* et produit un arbre de similarité entre les séquences. L'alignement des séquences est ensuite effectué en suivant les branches de cet arbre. Contrairement à BLAST, MUSCLE n'estime pas directement la probabilité qu'un alignement donné soit dû au hasard.

Bien que produisant des solutions optimales, les algorithmes de Needleman-Wunsch et de Smith-Waterman présentent des complexités trop élevées pour être applicables à l'analyse de près de 7 millions de protéines avec presque 48.000 protéines-requête. Les procédures itératives, bien qu'intéressantes dans la recherche d'homologues éloignés (entre chromosomes et plasmides) présentent deux inconvénients : i) elles convergent moins rapidement, et ii) dans le cas de familles multigéniques (par exemple, les recombinases/intégrases), le feront difficilement et identifieront l'ensemble des protéines codées par la famille multigénique pour une protéine-requête proche de cette famille [Guglielmini et al., 2013]. Les méthodes de type HMMER, bien qu'offrant de meilleures performances, dépendent grandement de la qualité des alignements de séquences fournis en entrée. La construction des alignements pour les différents groupes de protéines homologues identifiées constituerait une source supplémentaire de biais. Certains jeux de protéines (les protéines Xer notamment) produisent difficilement des alignements significatifs et une analyse de type HMMER reposant sur un alignement non-significatif constitueront un *profile* source d'erreur [Wistrand and Sonnhammer, 2005].

Le logiciel blastp de la suite BLAST+ [Camacho et al., 2009] a été choisi pour sa rapidité et son utilisation fréquente dans la communauté pour le calcul de similarité de protéines. Le seuil de rejet est fixé pour une e_{value} (probabilité d'obtenir un alignement significatif par hasard sachant la base de données et la protéine de référence) inférieure à 10^{-5} . Les scores de BLAST sont calculés par rapport au nombre r d'alignements locaux détectés (hits) de façon heuristique lors de la recherche de k-mers [Korf et al., 2003]. Le calcul de la e_{value} par BLAST peut être résumé par :

$$e_{value} = -ln(1 - p_{value})$$
où :
$$p_{value} = \frac{p'_{value}}{\beta^{r-1}(1 - \beta)}$$
avec :
$$p'_{value} = \frac{e^{-S_{sum}}S_{sum}^{r-1}}{r!(r - 1)!}$$
(4.1)

et où β est une constante fixée. S_{sum} est une fonction de la somme des scores S_i des hits détectés significatifs et de différentes variables telles que la taille T de l'espace de recherche selon l'algorithme utilisé [Korf et al., 2003] :

$$S_{sum} = K \sum_{i=1}^{r} S_i + f(r, T)$$
(4.2)

avec K constante. Les S_i , scores des hits d'un alignement de deux séquences $s_1^{S_i}$ et $s_2^{S_i}$ de longueur n peuvent être estimés par une équation de la forme [Korf et al., 2003] :

$$S_{i} = \sum_{j=1}^{n} ln \left(\frac{Q_{s_{1}^{S_{i}}, s_{2}^{S_{i}}}[j]}{P(s_{1}^{S_{i}}[j]) \cdot P(s_{2}^{S_{i}}[j])} \right)$$
(4.3)

où $P(s_1^{S_i}[j])$ et $P(s_2^{S_i}[j])$ sont les fréquences d'occurrences du *j*-ème caractère dans l'espace de recherche de $s_1^{S_i}$ et de $s_2^{S_i}$ respectivement et $Q_{s_1^{S_i}, s_2^{S_i}}[j]$, la fréquence d'occurrences de la paire formée par le *j*-ème caractère de $s_1^{S_i}$ et de $s_2^{S_i}$ dans l'espace de recherche.

Certains travaux ont étudié le lien entre score et pertinence biologique des différents algorithmes de recherche de similarité de séquence [Eddy, 2011]. En particulier, il a été mesuré qu'une e_{value} inférieure à 10^{-3} entre deux protéines correspond à 99% à des homologues fonctionnels (selon Pfam et leurs annotations de clan et pour une base de données de taille 192.987) [Boekhorst and Snel, 2007]. Un **cutoff** à 10^{-5} a été choisi et semble garantir un maximum de pertinence dans la recherche d'homologies fonctionnelles entre

les protéines-requête et les protéines identifiées. Un des biais peut cependant provenir de protéines contenant des domaines fonctionnels multiples et pouvant présenter des e_{value} très faibles avec des protéines partageant un de leurs domaines quoique possédant des fonctions totalement différentes [Song et al., 2007].

L'analyse des **6.903.452** protéines par les **47.604** protéines reliées aux STIG par *blastp* en utilisant une e_{value} seuil de 10^{-5} et les paramètres par défaut ([Camacho et al., 2009] a identifié **358.624 protéines homologues**, dont l'ensemble est désigné P_{homo} . La e_{value} d'une protéine p par rapport à une protéine q sera notée : $e_{value}(p,q)$.

4.2 Réalisation de clusters d'homologues protéiques et fonctionnels

Les protéines liées fonctionnellement aux STIG récupérées sont ensuite utilisées comme référence pour identifier des homologues fonctionnels dans les 5.125 réplicons de notre jeu de données. Ensuite, les protéines identifiées sont partitionnées selon leurs homologies de séquences *via* une analyse de clustering **afin de créer des unités d'homologies structurales et fonctionnelles**. Enfin, une procédure de "nettoyage" est effectuée afin d'exclure les clusters les plus biaisés (Figure 4.3).



FIGURE 4.3: Procédure de clustering des protéines.

Le choix et le paramétrage d'un algorithme de calcul d'homologie de séquences est suivi du choix et du paramétrage d'un algorithme de clustering de séquences protéiques. Différentes méthodes d'évaluation des clusters sont introduites.

4.2.1 Clustering des protéines

4.2.1.1 Définition

La majorité des algorithmes de clustering de séquences protéiques (ou nucléotidiques) reposent sur le calcul préalable d'une matrice de comparaison (aussi appelée matrice de distance ou de dissimilarité) de séquences deux à deux :

• Pour un ensemble de *n* protéines $P = \{p_1, ..., p_n\}$ et une distance *d* calculant un degré de dissimilarité de séquence $d(p_i, p_j)$ entre deux protéines p_i et p_j , soit M^d la matrice de distance où :

$$M_{ij}^d = d(p_i, p_j) \tag{4.4}$$

• S^d , la matrice de similarité, est alors définie par :

$$S_{ij}^{d} = 1 - \frac{d(p_i, p_j)}{d_{max}}$$
(4.5)

• Les graphes d'Interaction Protéines-Protéines (IPP) sont des graphes où S^d est la matrice d'adjacence du graphe $G = (P, E, S^d)$, où E est l'ensemble des couples (p_i, p_j) pour lesquels une homologie significative a été detectée.

4.2.1.2 Principe

Les algorithmes de calcul de similarité de séquences (Table 4.2) sont généralement utilisés pour le calcul de e_{value} qui servent de distances inter-protéiques. La majorité des méthodes de clustering de protéines sont en fait des algorithmes de détection de communautés appliqués aux IPP [Brohée and van Helden, 2006; Li et al., 2010]. Il existe de plus des critères externes spécifiquement développés pour les IPP. Un des biais potentiels dans l'utilisation des e_{value} des algorithmes en tant que distances est qu'elles ne sont pas obligatoirement **métriques**. En effet, *blastp* et *phmmer*, par exemple, produisent des e_{value} telles que $d(p_i, p_j) \neq d(p_j, p_i)$.

4.2.1.3 TRIBE-MCL

L'algorithme de clustering **TRIBE-MCL** [Enright et al., 2002] a été choisi pour son efficacité, démontrée notamment dans la construction des familles ACLAME [Leplae et al., 2010], pour sa capacité à regrouper des protéines multi-domaines [Enright et al., 2002; Frech and Chen, 2010] et pour ses meilleures performances dans l'identification de familles protéiques [Apeltsin et al., 2011; Frech and Chen, 2010].

TRIBE-MCL est dérivé de l'algorithme Markov Cluster (MCL) [van Dongen, 2000] qui est une méthode de détection de communautés. Le principe de l'algorithme est de capturer des clusters de protéines, régions du graphe où les protéines sont relativement plus interconnectées, en simulant des trajets aléatoires dans le graphe selon l'hypothèse qu'un trajet aléatoire dans un cluster tendra davantage à rester dans ce cluster que d'en sortir. Ces trajets aléatoires sont simulés *via* des opérations sur la matrice d'adjacence du graphe par deux opérations matricielles : *expansion* et *inflation* [Enright et al., 2002]. L'expansion a pour effet de dissiper les trajets aléatoires au sein des clusters, et l'inflation élimine les trajets inter-clusters. L'opération d'inflation est sous le contrôle d'un opérateur gr qui, pour des valeurs élevées, augmente le "resserrement" (ou **granularité**) des clusters obtenus. Une valeur de gr supérieure à 1 permet théoriquement la création de clusters. TRIBE-MCL prend comme mesure de similarité entre deux protéines p_i et p_j :

$$d(p_i, p_j) = -log_{10}(e_{value}(p_i, p_j))$$
(4.6)

L'efficacité de TRIBE-MCL à reconstruire des familles protéiques dépend grandement de la granularité choisie et sera influencée par la nature des familles de protéines étudiées [Apeltsin et al., 2011; Frech and Chen, 2010]. Ainsi, la valeur optimale de gr dépend du jeu de protéines analysé et de la question qu'on se pose.

4.2.2 Identification des domaines fonctionnels des protéines

Afin d'évaluer les clusters formés, les protéines ont été caractérisées selon leurs domaines fonctionnels. Pour l'identification des domaines des 358.624 homologues et des 47.604 protéines de référence, le programme hmmscan de la suite HMMER a été utilisé avec la base de données Pfam du 03/03/2013 comme base de données de référence, et, comme valeurs seuils, $e_{value} < 10^{-5}$ et $ce_{value} < 10^{-5}$ (conditionnal ce_{value}). La e_{value} pour un domaine est similaire à la e_{value} précédemment définie. La ce_{value} d'un domaine désigne la probabilité que le domaine ne soit pas significatif sachant que la protéine sur laquelle est trouvé le domaine est un vrai homologue d'une protéine de la base de données requête [Finn et al., 2011]. Les valeurs des autres paramètres sont celles par défaut [Eddy et al., 2013]. L'ensemble des domaines identifiés dans un ensemble de protéines Pest noté D_P . De même, l'ensemble des domaines identifiés sur les 358.624 protéines de $D_{\{p\}}$. **1.175.018** et **127.853** profiles ont ainsi été identifiés sur les 358.624 protéines de P_{homo} (ensemble des protéines homologues) et les 47.604 protéines de P_{ref} (ensemble des protéines de référence), respectivement. **1711** profiles différents sont identifiés (Figure 4.4), montrant l'hétérogénéité des protéines présentes.

4.2.3 Critères d'évaluation des clusters

L'évaluation des IPP se fait généralement sur des critères externes en les comparant à des familles ou classes de protéines de référence [Nepusz et al., 2012].

- ▶ Il est possible d'utiliser des indices dérivés de la sensibilité/spécificité et du F_{score} (éq. 3.9) [Brohée and van Helden, 2006].
- ▶ D'autres chercheurs proposent d'évaluer la p_{value} d'un cluster par rapport à une classe de référence selon leurs tailles [Li et al., 2010]. Pour un cluster C contenant kprotéines appartenant à une classe de référence F, la probabilité que ce cluster ne soit pas formé par hasard selon F est égale à :

$$p_{value} = 1 - \sum_{i=1}^{i=k-1} \frac{\binom{|F|}{i} \binom{|P|-|F|}{|C|-i}}{\binom{|P|}{|C|}}$$
(4.7)

► Un autre indice important dans la selection de la granularité pour les IPP est l'intracluster clustering coefficient (ICCC) [Lima-Mendez et al., 2008; Van Houdt et al., 2012]. Soit une protéine p de P, appartenant à un cluster C, et ayant un ensemble



FIGURE 4.4: Distribution des types et nombres d'occurence des profiles dans P_{homo} . 4.4a : Histogramme de l'occurrence des 1711 domaines Pfam identifiés. 4.4b : Distribution des nombres de domaines Pfam par protéine.
N de protéines voisines incluses dans C, pour tout i, j tels que $p_i, p_j \in N$, le clustering coefficient CC_p de p est défini par :

$$CC_p = \frac{\sum_{i=1}^{|N|-1} \sum_{j=i+1}^{|N|} S_{i,j}^d}{|N| \cdot (|N|-1)/2}$$
(4.8)

L'ICCC est alors défini par :

$$ICCC = \sum_{p \in P} CC_p \tag{4.9}$$

► Différents indices ont été utilisés pour répondre aux deux objectifs suivants : i) obtenir des clusters suffisamment homogènes en homologues fonctionnels et, ii) analyser les similitudes/dissimilitudes des réplicons via les clusters obtenus. Deux types de références sont identifiables dans les jeux de données : i) les familles de protéines ACLAME et KEGG et leur annotations, et ii) les domaines Pfam identifiés.

Pour tout $p \in P_{homo}$ et tout $p' \in P_{ref}$, une première annotation de p est donnée par :

$$Ann(p) = Ann(p') \iff e_{value}(p, p') = min\{e_{value}(p, p'') | p'' \in P_{ref}\}$$
(4.10)

L'objectif est d'estimer l'homogénéité en terme d'annotation des différents clusters mais l'*exhaustivité* des clusters n'est pas recherchée. Pour cela un indice externe de comparaison dérivé du Biological Homogeneity Index (BHI)
[Datta and Datta, 2006] est alors calculé. Soit un clustering Cl composé de k clusters : Cl = {C₁, ..., C_k}. Le BHI est alors défini par :

$$BHI = \frac{1}{|Cl|} \sum_{i=1}^{|Cl|} c_i \tag{4.11}$$

où c_i est défini par :

$$c_i = \frac{2}{|C_i| \cdot (|C_i| - 1)} \sum_{p, q \in C_i} d(p, q)$$
(4.12)

sous l'hypothèse que, pour tout $p \in C_i$, $\exists Ann(p)$ où d(p,q) est défini par :

$$d(p,q) = \begin{cases} 1 & \text{si } Ann(p) = Ann(q) \\ 0 & \text{sinon} \end{cases}$$

Contrairement à l'indice initial, la constante "2" est ajoutée pour que le dénominateur soit de la forme $\frac{N.(N-1)}{2}$, qui est le nombre de paires dans N éléments (on compare le nombre de *cas* observés sur le nombre de *cas* possibles, ainsi l'éq. 4.12, tout comme l'éq. 4.15 ci-après, est de la forme de l'éq. 4.8). En d'autres termes, le BHI est une mesure simple à interpréter, prenant ses valeurs dans [0, 1], qui est maximum si, pour tout cluster C_i de Cl, toutes protéines p de C_i possèdent une même annotation : Ann(p). Le BHI ne prend cependant pas en compte les potentielles grandes différences de taille entre

clusters (ce qui est notre cas : Figure 4.6). Une correction du BHI est alors apportée pour donner le *weighted BHI* (BHIw) :

$$BHIw = \frac{1}{|P|} \sum_{i=1}^{|Cl|} c_i \cdot |C_i|$$
(4.13)

où |P| est le nombre total de protéines du clustering (dans notre cas $|P_{homo}|$).

• Avec pour objectif de mesurer l'*homogénéité* des clusters, c'est-à-dire leur capacité à contenir des protéines ayant des domaines fonctionnels similaires, l'indice *Conservation Consistency Measure* (CCM) est introduit. Le CCM est défini par :

$$CCM = \frac{1}{|P|} \sum_{i=1}^{|Cl|} c'_i \cdot |C_i|$$
(4.14)

où c'_i est défini par :

$$c'_{i} = \frac{2}{|C_{i}| \cdot (|C_{i}| - 1)} \sum_{p,q \in C_{i}} d_{Jaccard}(D_{\{p\}}, D_{\{q\}})$$
(4.15)

et où $d_{Jaccard}$ est la distance de Jaccard pour deux ensembles. L'indice CCM prend ses valeurs dans [0, 1]. Une valeur proche de 0 indique des clusters qui contiennent des protéines ayant des domaines similaires.

• Enfin, on peut aussi calculer la proportion d'annotation majoritaire pour chaque cluster. Soient un cluster C, Ann_{max}^C l'annotation majoritaire des protéines de C, et $N_{Ann_{max}^C}$ le nombre de fois que $Ann(p) = Ann_{max}^C$ pour tout $p \in C$. On définit $Pr_{Ann_{max}}$ par :

$$Pr_{Ann_{max}} = \frac{N_{Ann_{max}}}{|C|} \tag{4.16}$$

4.2.4 Génération de clusters aléatoires

L'évaluation des clusters formés ainsi que les procédures de "cleaning" (ci-après) font intervenir des clusters de protéines engendrés aléatoirement. Une protéine étant uniquement caractérisée par le nombre et le type de domaines qu'elle comporte, on peut distinguer trois processus aléatoires indépendants intervenant dans la génération aléatoire d'un cluster de protéines :

- le nombre de protéines du cluster,
- le nombre de domaines fonctionnels présents dans une protéine donnée,
- le type d'un domaine donné.

4.2.4.1 Variables

Soit X_k , X_d et X_t les variables aléatoires discrètes et indépendantes correspondant aux nombre de protéines par cluster, nombre de domaines et type d'un domaine, respectivement. Les ensembles des valeurs possibles E_k , E_d et E_t sont alors déduits des données :

- $\mathbf{E}_{\mathbf{k}}$ prend ses valeurs entre 1 (le nombre minimum de protéines dans un cluster) et $max\{|C| \mid C \in Cl\}$ pour un Cl donné.
- $\mathbf{E}_{\mathbf{d}}$ prend ses valeurs entre 0 (le nombre minimum de domaines fonctionnels identifiés dans une protéine donnée) et $max\{|D_{\{p\}}| p \in P_{homo}\}$.

 $\mathbf{E}_{\mathbf{t}}$ prend ses valeurs dans $D_{P_{homo}}$.

De même, l'estimation des lois de probabilités de P_{X_k} , P_{X_d} et P_{X_t} se fait à partir des données :

$$\begin{split} \mathbf{P}_{\mathbf{X}_{\mathbf{k}}} & \text{Soient } O_{Cl} = \{ |C_i|, \ C_i \in Cl \} \text{ la distribution des tailles des clusters d'un clustering,} \\ & Q_{O_{Cl}}^q = \{Q_1, .., Q_q\} \text{ les q-quantiles de } O_{Cl}, \text{ et } x_{(p/q)} \text{ la première valeur de } Q_p \text{ (avec } x_{((q+1)/q)} = +\infty). \text{ On estime alors que :} \end{split}$$

$$P_{X_k} \to \begin{cases} P(x_{p/q} \le X_k < x_{(p+1)/q}) = \frac{|Q_p|}{|O_{Cl}|} & \text{où } p \in \{1, ..., q\} \\ P(X_k = x_{p_i}) = U([x_{p/q}, x_{(p+1)/q}])) & \text{où } x_{p_i} \in [x_{p/q}, x_{(p+1)/q}] \end{cases}$$
(4.17)

où U est la loi uniforme discrète. D'une part la probabilité d'obtenir une valeur de taille k incluse dans un certain intervalle dépend du nombre d'observations présentes dans cet intervalle et, d'autre part, toutes les valeurs de k au sein d'un intervalle donné ont la même probabilité d'être tirées.

 $\mathbf{P}_{\mathbf{X}_{\mathbf{d}}}$ En suivant la même démarche que pour P_{X_k} , on estime que :

$$P_{X_d} \to \begin{cases} P(x_{p/q} \le X_d < x_{(p+1)/q}) = \frac{|Q_p|}{|O_{P_{homo}}|} & \text{où } p \in \{1, ..., q\} \\ P(X_k = x_{p_i}) = U([x_{p/q}, x_{(p+1)/q}]) & \text{où } x_{p_i} \in [x_{p/q}, x_{(p+1)/q}] \end{cases}$$
(4.18)

avec $O_{P_{homo}}$ la distribution du nombre de domaines par protéine des protéines de P_{homo} et avec $Q_p \in Q^q_{O_{P_{homo}}}$.

 $\mathbf{P}_{\mathbf{X}_{t}}$ Soit d un domaine de D. On note N_{d}^{D} le nombre d'occurrences de d dans D. $P_{X_{t}}$ est alors simplement estimé par :

$$P(X_t = d) = \frac{N_d^{D_{P_{homo}}}}{|D|}$$
(4.19)

La création d'un cluster aléatoire s'effectue par un premier tirage avec remise sur X_d donnant la taille x_k du cluster. On effectue alors x_k tirages avec remise sur X_d et, pour chaque valeur x_d obtenue, on effectue finalement x_d tirages avec remise sur X_t .

4.2.4.2 Loi de probabilité des clusters aléatoires

Les variables aléatoires X_k , X_d et X_t étant indépendantes, soient X_C et X_{Cl} les variables indépendantes discrètes correspondant aux tirages d'un cluster aléatoire et d'un clustering aléatoire, respectivement. Soient $C_{k,P}$, un cluster aléatoire constitué d'un ensemble de k protéines $P = \{p_1, ..., p_k\}$, et $Cl_{rand} = \{C_{k_1,P_1}, ..., C_{k_z,P_z}\}$, un clustering aléatoire de z clusters. On peut alors estimer que :

$$P(X_C = C_{k,P}) = P(X_k = k) \cdot \prod_{p_i \in P} P(X_d = |D_{p_i}|) \cdot \prod_{d_j \in D_{p_i}} P(X_t = d_j)$$
(4.20)

et que :

$$P(X_{Cl} = Cl_{rand}) = \prod_{C_{k_i, P_i} \in Cl_{rand}} P(X_C = C_{k_i, P_i})$$

$$(4.21)$$

4.2.4.3 Vérification par test d'hypothèses

Lorsqu'est calculé un clustering aléatoire Cl_{rand} par rapport à un clustering observé Cl, une procédure de vérification est effectuée afin d'estimer si $O_{Cl_{rand}}$ et O_{Cl} , les distributions de taille des clusters de Cl_{rand} et Cl, respectivement, suivent la même loi. Pour cela, un test d'indépendance du χ^2 est réalisé entre $Q^p_{O_{Cl_{rand}}}$ et $Q^p_{O_{Cl}}$. Pour une p_{value} du test supérieure à 0.05, on estime que les deux distributions ne sont pas statistiquement différentes.

4.2.5 Protocole analytique

Pour chaque protéine de P_{homo} , une analyse blastp est conduite sur P_{homo} avec une valeur seuil de e_{value} de 10^{-5} permettant l'obtention d'une matrice de similarité S^d , avec d les e_{value} données par blastp. L'algorithme TRIBE-MCL est ensuite lancé sur S^d en utilisant les granularités gr suivantes : (2, 3, 4, 5, 6, 7, 8), et les indices CCM, BHI et BHIw sont alors calculés pour les différents gr. La variance des scores des trois indices obtenus avec des clustering aléatoires (non représentée) est globalement très faible (de l'ordre de 10^{-5}). La génération des clusters aléatoires a été réalisée en utilisant 1000 et 20 q-quantiles pour O_{Cl} et $O_{D_{P_{homo}}}$, respectivement, avec Cl, les clusterings obtenus pour les différents gr. L'adéquation des distributions O_{Cl} et $O_{Cl_{rand}}$ par le test du χ^2 est estimée en utilisant 100 q-quantiles.

4.2.5.1 Choix d'un gr de travail

Les différents indices confirment l'efficacité de TRIBE-MCL pour la création de partitions pertinentes dans le cas de protéines homogènes en annotations et en contenu de domaines fonctionnels (Figure 4.5).



FIGURE 4.5: Influence de la granularité (gr) sur le clustering des protéines homologues des STIG.

A : valeurs de BHI (orange), BHIw (vert). B : CCM (orange). C : Nombre de clusters de taille > 1 (gris) et de taille = 1 (rose). Losanges : résultats pour les clusterings engendrés aléatoirement.

Une augmentation de gr semble accroître la performance de l'algorithme. Cependant, un gr trop important ($gr \ge 3$) amplifie drastiquement le nombre total de clusters ainsi que le nombre de clusters ne contenant qu'une seule protéine. Une inflation trop importante semble ainsi avoir pour effet d'empêcher la détection de l'homologie entre protéines ou groupes de protéines, si elle est trop faible. Majorer gr a donc pour effet de faire perdre de l'information entre protéines et entre réplicons, tout en renforçant la pertinence biologique des clusters (dans l'objectif d'avoir des clusters de protéines homologues au niveau fonctionnel). En imaginant des cas extrêmes, nous pouvons faire l'hypothèse qu'un gr très important ne conservera l'homologie qu'entre des protéines très proches, appartenant vraisemblablement à des individus de la même espèce, et donc n'apportera pas d'information pertinente quant à la comparaison des différents réplicons. À l'inverse, un gr trop faible aura pour tendance : i) de créer de fausses homologies entre certaines protéines et de produire des liens erronés entre réplicons, et ii) de créer des clusters de protéines non pertinents d'un point de vue biologique. Un gr de 4 apparaît ainsi

être une valeur de travail pertinente, pour laquelle les valeurs de CCM et BHIw sont améliorées (Figure 4.5).

Un nombre raisonnable de clusters est produit tout en étant assez restrictif pour espérer séparer les groupes de protéines issues de familles multigéniques (Figure 4.6). La distribution en tailles k des clusters (Figure 4.6) ne semble pas suivre de lois de probabilité classique.



FIGURE 4.6: Distribution de la taille des clusters de protéines pour un gr de 4. Abscisse : nombre de protéines par cluster. Ordonnée : nombre d'occurrence.

Malgré une très grande majorité de clusters homogènes obtenus pour un Cl avec un gr de 4 (Figure 4.7), un faible pourcentage de clusters sont non-homogènes et regroupent des protéines ayant des annotations de fonctions proches (Table 4.3). Ce "bruit" semble être principalement lié à des familles multigéniques.

TABLE 4.3: Principales annotations multiples identifiées parmi les clusters de protéines des STIG et pourcentage de ces annotations sur l'ensemble des clusters hétérogènes.

Annotations multiples		
Ambiguïté XerC/XerD/Autres Xer	38	
Chromosomiques vs. plasmidiques ParA	11	
Chromosomiques vs. plasmidiques ParB	6	
Protéines DnaAB ambiguës	5	
Protéines Fts ambigües	5	



FIGURE 4.7: Pourcentage de l'annotation la plus fréquente par cluster, parmi l'ensemble des clusters (4.7a) et parmi les clusters ayant des annotations multiples (4.7b). Pourcentages calculés d'après l'éq. 4.16.

4.2.6 "Nettoyage" des clusters protéiques

En raison de la présence de protéines avec des domaines fonctionnels multiples, un seuil de 10^{-5} pour l'analyse blastp n'est pas suffisamment stringent pour garantir que la relation 4.10 est vérifiée dans tous les cas. Deux protéines partageant un même domaine fonctionnel peuvent être identifiées comme des homologues par une analyse de type blastp bien que ne dérivant pas forcément d'une même protéine ancestrale [Song et al., 2007]. Même si la question ici n'est pas de savoir si deux protéines ont une origine commune mais plutôt la même fonction, le problème reste identique : deux protéines peuvent posséder un domaine similaire (typiquement de type "transporteur ATP binding cassette", identifiant Pfam : ABC_tran, par exemple) en addition d'un autre domaine, et cependant avoir des fonctions totalement différentes. Concrètement, on retrouve ce problème dans P_{homo} avec, par exemple, l'obtention d'un cluster de protéines possèdant majoritairement un unique domaine de type "Sigma 54 modulation protein" (identifiant Pfam : Ribosomal_S30AE). Ces protéines sont toutes liées à une protéine annotées XerC de Porphyromonas (GI:332299940) qui comporte en plus d'un domaine de type intégrase, un domaine de type ABC_tran. Les protéines de ce cluster n'ont donc vraisemblablement pas le rôle d'intégrase et leur annotation XerC n'est donc pas pertinente.

4.2.6.1 Procédure de nettoyage

Pour un cluster C, on définit son annotation Ann(C) par :

$$Ann(C) = Ann(p) \iff N^{C}_{Ann(p)} = max\{N^{C}_{Ann(p_{i})} | p_{i} \in C\}$$
(4.22)

où $N_{Ann(p_i)}^C$ désigne le nombre de fois que $Ann(p_i)$ est trouvé pour les protéines de C. Soit une protéine $p \in P$ avec $P = P_{ref} \cup P_{homo}$ et $D_{\{p\}}$ son ensemble de domaines fonctionnels. Le vecteur des domaines de $p, v_p^{D_P}$, est alors introduit et est défini par :

$$v_p^{D_P} = (N_{d_1}^{D_{\{p\}}}, ..., N_{d_{|D_P|}}^{D_{\{p\}}}), \ d \in D_P$$
(4.23)

où $N_{d_i}^{D_{\{p\}}}$ est le nombre d'occurrences de d_i dans $D_{\{p\}}$. Pour un cluster de protéines C, on peut définir son vecteur de domaines $v_C^{D_P}$ par :

$$v_C^{D_P} = (\bar{N}_{d_1}^C, ..., \bar{N}_{d_{|D_P|}}^C)$$
(4.24)

où $\bar{N}_{d_i}^C$ est défini par :

$$\bar{N}_{d_i}^C = \frac{1}{|C|} \sum_{p \in C} N_{d_i}^{D_{\{p\}}}$$
(4.25)

Soit le clustering Cl_{ref} formé des différentes groupes d'orthologues KEGG et familles ACLAME tel que $Cl_{ref} = Cl_{KEGG} \cup Cl_{ACLAME}$. Pour un clustering Cl et pour tout $C_i, C_k \in Cl$, on considère alors la distance :

$$d_{eval}(C_i, C_k) = d_{cosine}(v_{C_i}^{D_P}, v_{C_j}^{D_P}) \text{ avec } C_j \in Cl_{ref} \text{ et } Ann(C_k) = Ann(C_j)$$
(4.26)

avec d_{cosine} , la distance *cosine*. Cette distance a une valeur unique pour chaque C_i car il n'existe qu'un seul $C_j \in Cl_{ref}$ tel que $Ann(C_k) = Ann(C_j)$. Soient un cluster $C_i \in Cl$ et un cluster aléatoire C_r tels que $|C_i| = |C_r|$. Soit X_{evalC_i} , la variable aléatoire de $d_{eval}(C_r, C_i)$ prenant ses valeurs dans $[0, +\infty[$. On considère alors que C_i est un cluster de Cl valide si et seulement si :

$$\begin{cases} d_{eval}(C_r, C_i) \leqslant x_{seuil_i} \\ P(x_{seuil_i} \leqslant X_{evalC_i}) = 0.90 \end{cases}$$

$$(4.27)$$

4.2.6.2 Estimation de x_{seuil_i}

Pour un cluster $C_i \in Cl$ donné, n tirages avec remise de clusters aléatoires C_{r_i} sont effectués avec $|C_i| = |C_{r_i}|$ en utilisant les estimations de X_d et X_t (relations 4.18 et 4.19). Soient Cl_{r_i} l'ensemble de ces clusters, $O_{Cr_i} = \{d_{eval}(C_{r_{i,1}}, C_i), ..., d_{eval}(C_{r_{i,n}}, C_i)\}$ l'ensemble des scores obtenus, et $Q_{O_{Cr_i}}^{10}$ ses déciles. On estime alors que :

$$x_{seuil_i} = x_{2/10} \tag{4.28}$$

où $x_{2/10}$ correspond à la première valeur du deuxième décile de $Q_{O_{C_r}}^{10}$.

4.2.6.3 Résultats et discussion

Cette procédure de cleaning consiste à évaluer si, pour un cluster de protéines C, les distributions en domaines fonctionnels de ses protéines se rapprochent de celles observées pour la famille de référence ayant la même annotation que C ($Ann(C) = Ann(C_{ref})|C_{ref} \in Cl_{ref}$). On évalue avec l'éq. 4.26 la distance séparant les deux clusters. Les deux clusters ont des protéines ayant des domaines similaires si la distance observée est inférieure à 90% des distances observées avec des clusters engendrés aléatoirement (éq. 4.27).

Le choix de la méthode de "cleaning", ses performances, ainsi que de possibles méthodes alternatives sont discutés plus loin. Cependant, un des biais à souligner est celui lié à $D_{P_{homo}}$, qui intervient dans l'estimation de X_t (éq. 4.19). À cause du jeu de protéines utilisé, certains domaines (tels que les domaines Pfam ABC_tran ou Phage_integrase) sont sur-représentés en comparaison à des domaines mineurs. Ainsi, des clusters aléatoires formés avec des protéines comportant ces domaines seront plus probables et la méthode d'évaluation pour, par exemple, un petit cluster de recombinases sera plus restrictive. De plus, des protéines fausses positives P_{FP} identifiées comme homologues à une protéine de référence multidomaines p_{ref} peuvent avoir une partie de leurs domaines en commun avec p_{ref} [Song et al., 2007], la procédure de cleaning ne prenant pas en compte ce biais dans la génération de clusters aléatoires. Enfin, un seuil de 10% étant assez large, il en résulte que la procédure de cleaning est finalement peu restrictive et n'élimine que les clusters comportant des protéines présentant des distributions de domaines très biaisées par rapport à leurs familles de références C_{ref} (Figure 4.8).

En utilisant un gr de 4, **917** clusters renfermant un total de **91.127** protéines, dont une large part est annotée "FtsE", ont été identifiés comme n'étant pas valides (Figure 4.8A). Ces protéines sont vraisemblablement des membres de la superfamille des transporteurs ABC. Une partie importante des clusters valides est aussi annotée FtsE (Figure 4.8B). Ces protéines possèdent généralement un unique domaine ABC_trans similairement aux "vraies" protéines FtsE. Il est alors probable que seules certaines d'entre elles ont une réelle fonction de type FtsE. On peut donc conclure que la seule utilisation de la distribution en domaines protéiques n'est pas toujours suffisamment discriminante pour l'identification des homologues fonctionnels. En considérant les seuils BLAST appliqués, on peut cependant supposer que ces protéines sont assez proches au niveau de leurs séquences et de leurs fonctions pour être tout de même maintenues dans notre jeu de données.



(A) Annotations des 917 clusters considérés comme incorrects.



(B) Annotations des 6465 clusters pertinents.

FIGURE 4.8: Annotations des clusters de protéines.

Les clusters sont annotés selon l'annotation majoritaire des protéines qu'ils renferment (éq. 4.10).

4.2.6.4 Alternatives à la procédure de cleaning

D'un point de vue méthodologique, la procédure de cleaning utilisée est similaire à une procédure de classification supervisée, effectuée pour chaque cluster de protéines obtenu C avec comme training set $E_{training} = \{E_{True}, E_{False}\}$. D'une part, l'ensemble C_{ref} des protéines a la même annotation que C (E_{True}) et, d'autre part, des clusters de protéines engendrés aléatoirement C_r selon la taille de C correspondent à la distribution du type et du nombre de domaines de l'ensemble des protéines récupérées par l'analyse par blastp (E_{False}). Différentes alternatives au calcul de la distance de l'éq. 4.27 peuvent alors être envisagées :

- Utiliser des algorithmes de classification supervisée classiques avec $E_{training} = \{C_{ref}, C_r\}$. Les protéines, en fonction de leurs domaines, sont alors classées comme étant similaires à C_{ref} ou à C_r et, en fonction du nombre de protéines identifiées positives, C est accepté ou refusé. Un des problèmes sous-jacents est de choisir les éventuels paramètres des algorithmes.
- Trier directement les protéines (et non les clusters) <u>avant</u> d'effectuer la procédure de clustering par TRIBE-MCL. Un des inconvénients de cette approche est la possibilité d'éliminer des protéines TP regroupées dans des clusters TP grâce à leur homologie de séquence et étant de vrais positifs bien que n'ayant pas de domaines fonctionnels identifiés. Cela est toutefois peu probable, /hmmscan étant beacoup plus sensible que /blastp.
- Modifier $E_{training}$ en incluant, par exemple, un ensemble plus large de protéines témoins TN dans E_{False} . Afin de compenser les biais introduits par des clusters FP identifiés par l'analyse à cause de la présence d'un domaine commun avec certaines protéines TP, E_{False} pourrait être constitué de clusters de protéines aléatoires ayant la même annotation que les protéines de E_{True} . Il est cependant possible que cette procédure rende l'étape de cleaning beaucoup trop discriminative, en particulier pour les petits clusters. Il aurait été également intéressant de tester C contre l'ensemble des E_{True} , et pas seulement contre l'ensemble présentant la même annotation que C, afin de limiter les ambiguïtés pour les annotations proches comme XerC/XerD.
- Inclure des données protéiques supplémentaires comme, par exemple, la taille et la synténie des domaines ou les données issues d'autres bases de données (TIGR-FAM...).
- Utiliser une méthode tirée de la littérature pour la discrimination fonctionnelle des protéines par leurs domaines. Song et. al [Song et al., 2008] proposent une méthode similaire reposant sur une classification par régression logistique et des pondérations différentes des domaines et des séquences. Cependant, des problèmes similaires dans la classification de protéines ayant des domaines fréquents (par exemple, pKinase) sont mis en avant. Enfin, différentes alternatives à MCL peuvent être prometteuses dans l'identification d'homologues [Terrapon et al., 2014].

Trouver la procédure optimisant le nombre de protéine TP présentes, homologues aux protéines de P_{homo} , et limitant le nombre de protéines FP demande des études additionnelles qui n'ont pas été la priorité de cette étude. La difficulté, ici, est que les protéines de P_{homo} appartiennent à différentes familles et présentent différents types d'homologies (Figure 4.9). Ainsi, un cluster peut être constitué de l'aggrégation de différents sousensembles de protéines homologues et, inversement, différents clusters peuvent appartenir à un même niveau d'homologie. La structuration des clusters est de plus fortement influencée par le choix des paramètres, ici, gr pour TRIBE-MCL, et dépend aussi des familles de protéines considérées.



FIGURE 4.9: Différents niveaux d'homologie chez les tyrosine-recombinases.

Chapitre 5

Séparation des réplicons

Les clusters de protéines obtenus sont ensuite utilisés pour caractériser les réplicons bactériens. L'objectif est de séparer les réplicons sur cette base et de voir si des regroupements spécifiques émergent. Les réplicons sont d'abord visualisés par des techniques de projection ou après transformation en graphe. Ensuite, ils sont clusterisés afin d'explorer plus spécifiquement les groupes formés et le positionnement des RECE par rapport aux autres réplicons.

5.1 Jeux de données obtenues et notations

5.1.1 Classification taxonomique

Les génomes bactériens sont organisés selon leur classification taxonomique, la hiérarchie des différents rangs taxonomiques suivant la relation :

$$\text{Espèce} \subset \text{Genre} \subset \text{Famille} \subset \text{Ordre} \subset Classe \subset Phylum \subset \text{Domaine}$$
(5.1)

Un individu (souche) d'une espèce bactérienne est caractérisé par son génome g, constitué d'un ou plusieurs réplicons. Soit $G = \{g_1, ..., g_{|G|}\}$, l'ensemble des génomes bactériens. On peut définir un génome g donné par l'ensemble des réplicons de différents types r_i qui le constituent :

$$g = \{r_1, \dots, r_{|g|}\}$$
(5.2)

5.1.2 Description des réplicons

On attribue aux réplicons le descripteur taxonomique de la bactérie qui les héberge. Ils possèdent de plus un type appartenant à l'ensemble {chromosome, plasmide, RECE} ("chr" est utilisé comme notation abrégée de "chromosome" dans les équations). Les clusters d'homologues de protéines des STIG permettent de caractériser les réplicons. Soient une solution de clustering des protéines $Cl = \{C_1, ..., C_{|Cl|}\}$, et l'ensemble des

réplicons $R = \{r_1, ..., r_{|R|}\}$. Un réplicon r peut être décrit par l'ensemble des gènes codant des protéines qu'il porte et est *de facto* défini par un ensemble de protéines :

$$r = \{p_1, ..., p_n\}$$
(5.3)

Un clustering de R, tel que $f_{clustering}(R) = \{\{r_{k_11}, ..., r_{k_1i}\}, ..., \{r_{k_x1}, ..., r_{k_xj}\}\}$, est noté Kl.

Un réplicon r est caractérisé par le vecteur v^r défini par :

$$v^r = (N^r_{C_1}, \dots, N^r_{C_{|C|}}) \tag{5.4}$$

où $N_{C_i}^r = |r \cap C_i|$ est le nombre de protéines de r dans C_i , un cluster de protéines homologues.

L'ensemble V^R des vecteurs v^r pour les réplicons r de R alors défini par :

$$V^{R} = \{v^{r_{1}}, ..., v^{r_{|R|}}\}$$
(5.5)

5.1.3 Groupes structuraux et taxonomiques

Soit une classe K de réplicons. On définit le vecteur pondéré des réplicons de K par :

$$\bar{v}^{K} = (\bar{N}_{C_{1}}^{K}, ..., \bar{N}_{C_{|Cl|}}^{K})$$
(5.6)

où $\bar{N}_{C_i}^K$ est défini par :

$$\bar{N}_{C_i}^K = \frac{1}{|K|} \sum_{r \in K} N_{C_i}^r$$
(5.7)

 \bar{v}^K est alors le centre de gravité de la classe K.

Les réplicons étant organisés par type et par groupe taxonomique, les notations suivantes sont introduites :

- R_{taxon} désigne l'ensemble des réplicons de R dont l'hôte appartient au groupe taxonomique taxon. $V^{R_{taxon}}$ désigne, par extension, l'ensemble des vecteurs v^r des réplicons r de R_{taxon} .
- $R^{\{type\}}$ désigne l'ensemble des réplicons, $r \in R$, où $Type(r) \in \{type\}$ et où Type(r) est la fonction renvoyant le type de r.
- Pour un réplicon r donné, la fonction Tax_{n_tax}(r) renvoie la valeur du rang taxonomique considéré de r avec n_tax appartenant à la relation 5.1. Par exemple, Tax_{phylum}(r) = Protéobactéries ou Tax_{class}(r) = Bêta-protéobactéries.
- $Kl_{n.tax}^R = \{K_1, ..., K_{|Kl|}\}$ est l'ensemble des classes K_i de réplicons où $K_i \subset R$ tel que :

$$\begin{cases} Tax_{n_tax}(r_x) = Tax_{n_tax}(r_y) & r_x, r_y \in K_i \\ Tax_{n_tax}(r_x) \neq Tax_{n_tax}(r_z) & r_z \in K_{j, \ j \neq i} \end{cases}$$

Par exemple, Kl_{phylum}^R sera l'ensemble des réplicons de R organisés en classes selon le phylum d'appartenance de leur hôte.

L'ensemble $\bar{V}^{Kl_{n_tax}^R}$ des \bar{v}^{K_i} où $K_i \in Kl_{n_tax}^R$ est alors défini par :

$$\bar{V}^{Kl^R_{n_tax}} = \{ \bar{v}^{K_1}, ..., \bar{v}^{K_{|Kl^R_{n_tax}|}} \}$$
(5.8)

On définit $\bar{V}_{n_tax}^R$ comme l'ensemble des vecteurs de réplicons de R normés selon n_tax et regroupés par *type* tel que :

$$\bar{V}_{n_tax}^{R} = \bar{V}^{Kl_{n_tax}^{R\{chr\}}} \cup \bar{V}^{Kl_{n_tax}^{R\{plasmide\}}} \cup \bar{V}^{Kl_{n_tax}^{R\{RECE\}}}$$
(5.9)

Par extension $\bar{V}_{n,tax}^{R_{taxon}}$ désigne l'ensemble des vecteurs de réplicons de R_{taxon} normés selon n_{tax} et regroupés par type.

Pour un ensemble $Kl_{n_tax}^{R_i}$ donné avec $R_i \subset R$, et pour $K \in Kl_{n_tax}^{R_i}$, si tous les réplicons de K sont du même type, on peut annoter K selon les valeurs de $Tax_{n_tax}(r)$ et de Type(r) où $r \in K$. De même que pour un réplicon, on peut alors définir $Tax_{n_tax}(K)$ et Type(K) tels que $Tax_{n_tax}(K) = Tax_{n_tax}(r)$ et Type(K) = Type(r) pour $r \in K$. On définit $\bar{Kl}_{n_tax}^{R_i}$, l'ensemble des couples ($Type, Tax_{n_tax}$) existant dans R_i par :

$$\bar{K}l_{n_tax}^{R_i} = \left\{ \{Type(K), Tax_{n_tax}(K)\} \mid K \in Kl_{n_tax}^{R_i} \right\}$$
(5.10)

 $\bar{K}l_{n_tax}$, l'ensemble des couples (*Type*, Tax_{n_tax}) existant dans R est alors défini par :

$$\bar{K}l_{n_tax} = \bar{K}l_{n_tax}^{R\{chr\}} \cup \bar{K}l_{n_tax}^{R\{plasmide\}} \cup \bar{K}l_{n_tax}^{R\{RECE\}}$$
(5.11)

5.1.4 Dimension des données

Sur l'ensemble des 5125 réplicons considérés et sur les **267.497** homologues identifiés (358.624 homologues présumés diminués des 91.127 faux positifs) pour une granularité gr de 4 pour TRIBE-MCL, **4928** réplicons codent au moins une protéine parmi les homologues. **6096** clusters de protéines sont formés et sont considérés comme authentiques (ils ne correspondent pas à de faux positifs).

TABLE 5.1: Dimension des données utilisées.

R est l'ensemble des réplicons ayant au moins une proté ine homologue. Cl est la solution de clustering obtenue pour un gr de 4 après la procédure de cleaning.

Données	Matrice	Taille
R	-	4928
Cl	-	6096
V^R	M^{V^R}	(4928, 6096)
\bar{V}_{genre}^R	$M^{\bar{V}^R_{genre}}$	(851, 6096)
\bar{V}^R_{classe}	$M^{\bar{V}^R_{classe}}$	$(91,\!6096)$
$ar{V}^R_{phylum}$	$M^{\bar{V}^R_{phylum}}$	(60, 6096)

5.2 Méthodes d'évaluation de la séparation des réplicons

L'évaluation des méthodes de clustering est fondée sur des critères internes et/ou externes. Dans le cas de critères externes, la difficulté est de choisir les classes de référence auxquelles les clusters obtenus seront comparés. Ici, les clusters de réplicons obtenus sont analysés des points de vue structural (selon le type de réplicons) et taxonomique. L'efficacité des différentes méthodes de visualisation des données peut être estimée visuellement en inspectant la capacité des méthodes à séparer les données selon les critères utilisés pour le clustering : taxonomie et type, ainsi qu'en utilisant une procédure de clustering sur les données projetées suivie de la mesure de critères externes.

5.2.1 Critères de validation externe utilisés

La structure génomique des STIG chromosomiques et plasmidiques différant (Chapitre 1), une étude des réplicons bactériens selon leurs STIG doit mener à une discrimination nette des chromosomes et plasmides en des catégories distinctes. On peut postuler que les gènes liés aux STIG font, dans leur ensemble, partie du *génome-cœur* des chromosomes. Il en ressort deux prédictions quant à la classification des réplicons par leurs STIG : i) les génomes stables (*i.e.*, chromosomes) des individus de la même espèce auront des STIG très proches, et ii) l'homologie inter-espèces des STIG chromosomiques est retrouvée pour des représentants d'espèces proches phylogéniquement. Au final, deux critères externes sont mesurés :

- la séparation et l'homogénéité des clusters de réplicons selon leur type,
- l'homogénéité des clusters de réplicons selon la classification taxonomique de leur hôte.

La comparaison externe de deux clusterings s'effectuera sur la base des scores obtenus pour ces deux critères. Un clustering pertinent, décrivant les réalités génomiques, aura une grande probabilité de respecter ces deux critères. L'hypothèse sur le second critère ne concernera que les réplicons chromosomiques et aucune supposition n'est faite quant aux scores obtenus pour des clusters de plasmides.

Les données utilisés sont V^R avec R, l'ensemble des 5125 réplicons. À cause de la disparité de représentation de certains genres bactériens (par exemple, *Escherichia, Vibrio, Burkholderia...* sont représentés par de nombreuses espèces et/ou souches, alors qu'un seul génome d'une seule espèce de *Paracoccus, Asticaccaulis, Sphaerobacter* n'est présente dans V^R), l'ensemble \bar{V}_{genre}^R est utilisé au lieu de V^R dans certains cas. \bar{V}_{genre}^R permet d'étudier non plus les individus mais les genres bactériens. Pour chaque couple $\{genre, type\}$, un vecteur normé est construit avec comme valeur de chaque variable la moyenne des valeurs pour les variables des réplicons identifiés par ce couple (éq. 5.9). Pour un clustering de réplicons Kl formé, différents ensembles de réplicons de référence Kl_{ref} sont alors formés selon le type des réplicons, la taxonomie des espèces hôtes des réplicons, et selon que les données sont normées ou pas (Table 5.2). L'évaluation des solutions de clustering des réplicons se fera par comparaison avec cinq ensembles de référence (Table 5.2).

TABLE 5.2: Ensemble	s des classes	de référence	Kl_{ref}	selon	la séparation	étudiée.
---------------------	---------------	--------------	------------	-------	---------------	----------

Séparation étudiée	Ensembles non normés	Ensembles normés
type des réplicons phylum des chromosomes phylum des plasmides classe des chromosomes classe des plasmides	$\{R^{\{plasmide\}}, R^{\{chr\}}\}$ Kl^{chr}_{phylum} $Kl^{plasmide}_{phylum}$ Kl^{chr}_{classe} $Kl^{plasmide}_{classe}$	$ \begin{aligned} & \{\bar{K}l_{genre}^{R\{plasmide\}}, \bar{K}l_{genre}^{R\{chr\}} \} \\ & \{\bar{K}l_{genre}^{K} \ K \in Kl_{phylum}^{chr} \} \\ & \{\bar{K}l_{genre}^{K} \ K \in Kl_{phylum}^{plasmide} \} \\ & \{\bar{K}l_{genre}^{K} \ K \in Kl_{classe}^{chr} \} \\ & \{\bar{K}l_{genre}^{K} \ K \in Kl_{classe}^{chr} \} \end{aligned} $

5.2.1.1 V-measure

La V-measure [Rosenberg and Hirschberg, 2007] est la moyenne harmonique des deux indices homogeneity et completeness, qui traduisent, respectivement, les degrés d'homogénéité et d'exhaustivité des clusters (Chapitre 3). Soient un ensemble d'observations $E = \{e_1, ..., e_{|E|}\}, Kl$ un clustering de E, et Kl_{ref} un ensemble de classes de référence des observations de E. La V-measure est définie par :

$$V - measure = 2. \frac{homogeneity.completeness}{homogeneity + completeness}$$
(5.12a)

où les indices homogeneity et completeness sont définis par :

$$homogeneity = \begin{cases} 1 & \text{si } H(Kl|Kl_{ref}) = 0\\ 1 - \frac{H(Kl|Kl_{ref})}{H(Kl)} & \text{sinon} \end{cases}$$
(5.12b)

$$completeness = \begin{cases} 1 & \text{si } H(Kl_{ref}|Kl) = 0\\ 1 - \frac{H(Kl_{ref}|Kl)}{H(Kl_{ref})} & \text{sinon} \end{cases}$$
(5.12c)

avec H(Kl), l'entropie de Kl, définie par :

$$H(Kl) = -\sum_{C \in Kl} \frac{|C|}{|E|} \cdot \log(\frac{|C|}{|E|})$$
(5.12d)

et $H(Kl|Kl_{ref})$, l'entropie conditionnelle de Kl sachant Kl_{ref} , telle que :

$$H(Kl|Kl_{ref}) = -\sum_{C \in Kl} \sum_{K \in Kl_{ref}} \frac{|C \cap K|}{|E|} . log(\frac{|C \cap K|}{|K|})$$
(5.12e)

Les indices homogeneity, completeness et V-measure sont, respectivement, des analogues des indices de précision, sensibilité et F-measure (éq. 3.9). Cependant, de par leurs définitions et contrairement aux indices de la F-measure, ils prennent en compte de façon plus cohérente la taille et l'hétérogénéité des clusters et présentent des meilleures performances dans l'évaluation de plusieurs clusterings [Rosenberg and Hirschberg, 2007]. Homogeneity et completeness prennent leurs valeurs dans l'intervalle [0, 1], une valeur

de 1 indiquant de parfaites homogénéité et exhaustivité de Kl par rapport à Kl_{ref} . Dans les évaluations des clusters des réplicons, seule l'homogeneity est vraiment discriminante. La completeness est un indicateur du niveau auquel des classes différentes sont réparties dans des clusters distincts.

5.2.2 Critères de validation interne

5.2.2.1 Coefficient silhouette

Dans un premier temps, nous avons mesuré les qualités intrinsèques des clusterings obtenus par des critères de performance classiques comparant les distances des observations intra-clusters et inter-clusters. Le **coefficient silhouette** (éq. 3.11) a d'abord été utilisé afin de mesurer l'enchevêtrement des clusters obtenus, notre objectif étant qu'il soit minimisé. Les mesures obtenues par ce critère semblent cependant peu pertinentes compte tenu de la très grande dimensionalité des données, et en comparaison des critères externes utilisés qui permettent d'évaluer le pouvoir de séparation d'un clustering en se référant à des concepts biologiques plus "concrets" que les distances mesurées par ce coefficient.

5.2.2.2 Critère de stabilité

Nous avons ensuite utilisé un critère de robustesse, estimant la stabilité des clusters. Un cluster stable témoignera soit d'une réalité analytique par rapport aux données utilisées, soit de l'inflexibilité d'une méthode particulière de clustering envers un jeu de données particulier [Hennig, 2007]. Le caractère instable d'un cluster reflètera toujours une probabilité d'existence faible de celui-ci qui alors ne devra pas être interprété [Hennig, 2007]. L'approche utilisée, inspirée de Hennig [Hennig, 2007], consiste, pour un clustering donné Kl, à effectuer des ré-échantillonages $Ech_{Kl} = \{Kl_1, ..., Kl_n\}$ pour évaluer la conservation des clusters C de Kl dans les différents Kl_i de Ech_{Kl} .

Soit un ensemble $E = \{e_1, ..., e_{|E|}\}$ d'observations, une procédure de clustering $f_{clustering}$, et un clustering Kl de E tel que $f_{clustering}(E) = Kl$. On applique la procédure suivante :

- (1) Un bootstrap sur E de n échantillons, $Ech_E = \{E_1, ..., E_n\}$, est réalisé à partir de tirages aléatoires avec remise sur E
- (2) Pour chaque échantillon E_i de Ech_E , on calcule $f_{Cl}(E_i)$. On note $Ech_{Kl} = \{f_{Cl}(E_i) | E_i \in Ech_E\}$
- (3) Soient un cluster C tel que $C \in Kl$, un clustering Kl_i tel que $Kl_i \in Ech_{Kl}$, et un ensemble d'observations E_i tel que $E_i \in Ech_E$. On définit \triangle_i^C , comme la valeur :

$$\Delta_i^C = \begin{cases} \max\{d_{Jaccard}(C, C_i^{ech}) \mid C_i^{ech} \in Kl_i\} & \text{si } C \cap E_i \neq \emptyset \\ -1 & \text{sinon} \end{cases}$$
(5.13)

(4) Soit $V_{\Delta_i^C}$, les valeurs prises par Δ_i^C . L'estimateur de stabilité du cluster C, Δ^C , est défini par :

$$\Delta^C = \frac{1}{|B_C|} \sum_{E_i \in Ech_E, \ \Delta_i^C \neq -1} \Delta_i^C \tag{5.14}$$

avec :

$$B_C = \{ \Delta_i^C \mid \Delta_i^C \in V_{\Delta_i^C} \text{ et } \Delta_i^C \neq -1 \}$$

(5) \triangle^{Kl} , l'estimateur de stabilité du clustering Kl, est alors défini par :

$$\Delta^{Kl} = \frac{|C|}{|E|} \sum_{C \in Kl} \Delta^C \tag{5.15}$$

(6) Soit $Cm_C^{Kl_i}$, un cluster de Kl_i tel que $\Delta_i^C \neq -1$ et que $d_{Jaccard}(C, Cm_C^{Kl_i})$ est minimale pour l'ensemble des clusters de Kl_i . Un **estimateur de stabilité d'une observation** e pour $e \in C$, Δ^e , est défini par :

$$\Delta^e = \frac{|B_C^e|}{|B_C|} \tag{5.16}$$

avec :

$$B_C^e = \{ e \in Cm_C^{Kl_i} | Kl_i \in Ref_{Kl} \}$$

L'indice \triangle^{Kl} est la moyenne des \triangle^{C} pondérée selon la taille des clusters de Kl, ce qui diffère de la procédure originale de Hennig. \triangle^{Kl} permet de comparer des solutions de clustering en donnant plus de poids aux clusters de grande taille. Pour un cluster C de Kl et une observation e de C, \triangle^{e} permet d'évaluer la stabilité de e dans C en évaluant le nombre de fois que e est retrouvé dans les clusters $Cm_{C}^{Kl_{i}}$ des Kl_{i} de Ech_{Kl} .

 \triangle^e, \triangle^C et \triangle^{Kl} prennent leurs valeurs dans l'intervalle [0, 1]. Une valeur de \triangle^C inférieure à 0.5 indique un cluster "dissous" et non interprétable, alors qu'une valeur supérieure à 0.75 indique un cluster stable [Hennig, 2007, 2008].

La méthode de Fang et Wang [Fang et al., 2010], produisant un indice de stabilité similaire, a aussi été employée pour la selection du choix du nombre de clusters à utiliser en *input* de l'algorithme WARD (voir ci-après).

5.2.3 Sélection de modèle

5.2.3.1 Critères

► Les choix des algorithmes de clustering, des méthodes de projections et des paramètres sont guidés par les scores d'homogeneity obtenus pour la séparation de l'ensemble

des réplicons selon leur type, et pour la séparation des chromosomes selon leur taxonomie (phylum et classe). De façon complémentaire, l'indice de stabilité \triangle^{Kl} des clusterings obtenus est également pris en compte.

- Pour estimer la capacité des algorithmes à produire des clusters uniques pour chaque unité taxonomique, l'indice de *completeness* est aussi considéré pour les séparations des chromosomes en fonction de leur appartenance à un phylum et à une classe taxonomique. Idéalement, une partition des données produit des clusters stables séparant plasmides et chromosomes et où, pour les clusters de chromosomes, la taxonomie des espèces-hôte des réplicons est retrouvrée.
- ▶ Pour un algorithme donné, la sélection des paramètres est principalement effectuée en fonction des scores de stabilité \triangle^{Kl} et de l'indice de stabilité de Fang et Wang [Fang et al., 2010]. En particulier, lorsque le nombre de clusters k doit être entré, la valeur de k est fixée lorsque une augmentation de k n'entraîne pas d'accroissement significatif des scores des indices de stabilité des clusterings obtenus.
- ▶ Les indices \triangle^C et \triangle^e permettent d'évaluer la pertinence d'un cluster ou d'une observation à l'intérieur d'un cluster donné.

5.2.3.2 Sélection de modèle pour les méthodes de projection

Dans un premier temps, les différentes représentations des données sont comparées visuellement en estimant leur efficacité à structurer les données selon le type de réplicon ou la taxonomie de l'espèce-hôte. Afin de calculer des indicateurs, les données obtenues par les projections, $f_P^2(V^R)$, sont soumises à une même procédure de clustering WARD utilisant un nombre de clusters fixé à 50. Les scores d'homogeneity, completeness, Vmeasure et de stabilité \triangle^{Kl} sont ensuite calculés à partir de ces clusters et permettent d'obtenir des mesures chiffrées pour les différentes projections de données. Cependant, si les projections ne produisent pas les mêmes types de partition, l'identification des partitions nécessitera des procédures spécifiques à chaque projection, ce qui risque d'apporter des biais à l'analyse. Nous estimons néanmoins que cette méthodologie "maison" permet de proposer une première classification de l'efficacité des différentes projections du jeu de données.

5.3 Visualisation des données

Cette étape vise à représenter spatialement l'organisation des réplicons selon leur contenu en gènes des STIG de façon interprétable et pertinente, et d'identifier le positionnement des RECE par rapport aux autres réplicons. Les études traditionnelles de génomique comparative ont souvent recours à des structures de type arbre pour représenter les liens entre différents éléments génétiques, les liens recherchés étant classiquement des liens phylogéniques et évolutifs. Ces approches ne sont pas appropriées pour représenter les relations entre réplicons pour plusieurs raisons. i) Compte tenu des diversités des réplicons et de leurs STIG, la comparaison de deux réplicons peut difficilement être transformée en distance évolutive. ii) Un modèle de descendance linéaire entre les réplicons est simpliste et ne prend pas en compte les phénomènes d'échanges latéraux à l'œuvre dans les génomes bactériens [Bapteste et al., 2009; Doolittle and Bapteste, 2007]. iii) Des mosaïques de STIG caractérisent les réplicons bactériens (*cf.* Chapitre 1). Étudier une famille donnée de gènes ou de protéines, par exemple, les ParA/ParB, n'est pas suffisant pour décrire la diversité des STIG existant au sein des réplicons, l'objectif sous-jacent étant précisément de rendre compte de cette diversité.

L'exploration des données a donc reposé sur deux approches méthodologiques de visualisation :

- La projection des données dans un espace de dimension réduite, 2D ou 3D, interprétable visuellement.
- La transformation des données en graphe (bipartite) et sa visualisation par "spatialisation".

5.3.1 Réduction de dimension des données par projection

5.3.1.1 Méthodes utilisées

Quatre types de projection ont été étudiés pour réaliser la projection des ensembles V^R et \bar{V}^R_{aenre} dans des espaces à deux dimensions (Table 5.4).

TABLE 5.4: Méthodes de projection utilisées dans l'analyse des données.

- **ACP** L'Analyse en Composantes Principales [Hotelling, 1933] est une méthode statistique permettant d'exprimer les valeurs de p variables corrélées tirées de n observations en q variables non-corrélées, de manière à ce que les q variables soient des combinaisons linéaires des p variables et qu'elles soient orthogonales entre elles selon une distance d. Les q variables, ou composantes, définissent un espace en q dimensions (q < p) dans lequel sont projetées les observations. Les q variables sont choisies pour que la variance entre observations soit maximisée (et donc que l'*inertie* des observations par rapport aux q composantes soit minimale [Duby and Robin, 2006]). Le seul paramètre à choisir est q, le nombre de dimensions, d étant la distance Euclidienne.
- **ISOMAP** Cette procédure est une méthode de réduction non-linéaire ne conservant que localement les distances entre une observation et les autres [Tenenbaum, 1998; Tenenbaum et al., 2000]. Pour une observation o et une distance d donnée, l'algorithme cherche les k voisins o_{k_i} de o tels que les $d(o, o_{k_i})$ soient minimales. Ces distances sont alors conservées. Pour calculer la distance séparant r d'une observation n'appartenant pas à ses k plus proches voisins, la distance géodésique correspondant au chemin le plus court dans le graphe formé par les k plus proches voisins est alors utilisée. Ces distances estimées sont ensuite utilisées dans une procédure de type MDS. Les paramètres de l'algorithme ISOMAP sont la distance d initialement utilisée (généralement la distance Euclidienne), le nombre de voisins k, et les paramètres liés à la procédure MDS.

MDS L'ensemble de procédures *MultiDimensional Scaling* permet de positionner des observations dans un espace de dimension réduite [Izenman, 2008]. Un exemple classique de procédure MDS métrique, f_{MDS}^q , pour un ensemble d'observations dont les vecteurs associés sont rassemblés dans V, consiste à minimiser la quantité :

$$Stress = \sum_{i < j} \left(d_{euclidienne}(v_i, v_j) - d(f_{MDS}^q(v_i), f_{MDS}^q(v_j)) \right)^2$$
(5.17)

pour tout v_i et tout v_j de V, et où $d_{euclidienne}$ est la distance euclidienne. Cette quantité peut servir à valider la projection des données (typiquement Stress < 0.20 [Izenman, 2008]). Les paramètres à choisir sont la distance d utilisée pour comparer les observations et des paramètres propres à la méthode de convergence de Stress vers un minimum (nombre d'itérations, d'instantiation...)

SOM

L'algorithme de Self-Organizing Map [Kohonen, 1982] consiste à assimiler chaque observation o et son vecteur associé v_o à l'une des cellules d'un ensemble de cellules D organisées en q dimensions. Plus précisément, une cellule c est représentée par un vecteur v_c de même dimension que v_o et est assimilée à o si et seulement si $d(v_o, v_c) = min\{d(v_o, v_c) | c \in D\}$. Ce processus est itératif et les valeurs de v_c sont modifiées à chaque itération pour tendre vers celles de v_o . Les coordonnées de c, $\{x_1, ..., x_q\}$ dans D, correspondent alors à $f_{SOM}^q(v_o)$ si o est assimilée à cà la dernière itération [Izenman, 2008]. Les paramètres à choisir sont le nombre d'itérations, les dimensions de D et la distance d utilisée.

La projection par ACP a d'abord été utilisée mais n'a pas permis d'obtenir une structuration pertinente des réplicons (voir ci-après). D'autres méthodes ont alors été employées. La procédure MDS a été testée avec la distance *cosine* (éq. 3.4) ainsi qu'avec la distance Euclidienne modifiée (éq. 3.3). L'algorithme SOM, utilisé dans des domaines d'étude variés, constitue une alternative originale aux méthodes de visualisation classiques. Nous avons également eu recours à l'algorithme ISOMAP qui a été créé spécialement pour prendre en compte la non-linéarité de certains jeux de données, ce qui justifie son choix dans l'analyse.

5.3.1.2 Logiciels utilisés

- Scikit-learn librairie de Machine learning de Python a été utilisée pour le calcul des projections par ACP, MDS et ISOMAP, et celui des indices de V-measure (homogeneity, completeness et V-measure) et des clusters WARD.
- **Python** a été utilisé pour réaliser les nombreux workflows, préparer les données, et calculer les distances inter-observations.

5.3.1.3 Paramètres des analyses

- ▶ Les choix du nombre d'itérations et des dimensions des cellules de SOM ont suivi les recommandations pour cette méthode [Wendel and Buttenfield, 2010].
- ▶ Pour l'algorithme MDS, le choix de la distance s'est porté sur la distance *cosine* (éq. 3.4), qui produit des résultats significativement meilleurs que les distances euclidiennes. Les paramètres additionnels sont les paramètres par défaut.

- ▶ Le nombre de voisins d'ISOMAP (k = 5) a été choisi en fonction des valeurs des indices d'*homogeneity*.
- ▶ Le nombre de clusters k utilisés par l'algorithme WARD sur les projections $M_P^{V^R}$ et $M_P^{\bar{V}_{genre}^R}$ de, respectivement, V^R et \bar{V}_{genre}^R a été choisi sur la base des scores de l'indice de stabilité de Fang et Wang obtenus. Un unique k est choisi pour des différentes projections $M_P^{V^R}$ pour une raison d'homogénéité.

5.3.1.4 Projection de V^R et de \bar{V}^R_{qenre}

Les méthodes de projection par ACP et ISOMAP se sont révélées inefficaces pour la représentation des réplicons. On peut noter des scores particulièrement faibles des indices d'*homogeneity* obtenus pour la structuration des chromosomes selon la classe ou le phylum de leur hôte par ces méthodes (Table 5.5). Ces méthodes sont globalement moins performantes, et très peu performantes à séparer les chromosomes selon l'appartenance de leur hôte à un groupe taxonomique.

	\mathbf{Indice}^{a}	SOM	MDS	\mathbf{PCA}^d	ISOMAP
$Paramètres^b$		<i>it</i> :60000	d: cosine		k :5
Nombre de clusters ^{c}		50	50	50	50
	homogeneity	0.84	0.85	0.79	0.71
Type de réplicons	completeness	0.18	0.18	0.25	0.19
	V-measure	0.29	0.29	0.38	0.30
	homogeneity	0.83	0.78	0.66	0.70
Phylum des chromosomes	completeness	0.50	0.46	0.33	0.35
	V-measure	0.62	0.58	0.44	0.47
	homogeneity	0.85	0.78	0.68	0.73
Classes des chromosomes	completeness	0.75	0.68	0.49	0.54
	V-measure	0.80	0.72	0.57	0.62
Phylum des plasmides	homogeneity	0.65	0.41	0.05	0.20
	completeness	0.31	0.26	0.12	0.17
	V-measure	0.42	0.36	0.07	0.18
	homogeneity	0.62	0.41	0.06	0.17
Classes des plasmides	completeness	0.41	0.27	0.22	0.21
	V-measure	0.50	0.33	0.10	0.19

TABLE 5.5: Évaluation des procédures de visualisation de l'ensemble V^R des réplicons.

 $^a~V\mbox{-}measure$ calculées selon l'éq. 5.12.

 b it, nombre d'itérations utilisées par SOM, d distance utilisée par l'algorithme MDS et k, nombre de voisins considérés par l'algorithme ISOMAP.

La difficulté qu'a l'ACP à séparer chromosomes et plasmides peut provenir du fait qu'ils ne sont pas linéairement séparables. La non-linéarité résulte en partie de la très grande

^c nombre de clusters donnés en *input* pour l'algorithme WARD.

d variance expliquée par les 2 principales composantes : 16% .

dimension des données. La distance euclidienne entre deux chromosomes proches (à un certain degré) peut être plus élevée que celle séparant deux plasmides n'ayant rien en commun (*cf.* Chapitre 3 §3.4.7.2). De nombreux réplicons, dont l'ensemble des plasmides, sont agrégés en un unique groupe (Figures 5.1a et 5.1b), témoignant de la difficulté des méthodes ACP et ISOMAP à structurer des observations qui ne possèdent que quelques attributs non nuls.



FIGURE 5.1: Projection de l'ensemble des réplicons V^R . Coloration selon le type de réplicon : chromosome (gris clair), plasmide (gris foncé), RECE (bleu).

Les algorithmes SOM et MDS produisent une organisation cohérente des réplicons bactériens (Table 5.5; Figures 5.1c et 5.1d). Ils fournissent une meilleure organisation taxonomique des chromosomes que les ACP et ISOMAP, et réussissent à structurer les

réplicons plasmidiques (Table 5.5). La distance *cosine*, utilisée par les deux méthodes SOM et MDS, ne tient compte que des attributs non-nuls des observations, et produit des distances pertinentes entre plasmides et chromosomes ainsi qu'entre plasmides. Cette structuration n'est de plus pas influencée par le biais d'échantillonnage des données car elle est également observée sur \bar{V}_{qenre}^R (Figures 5.2a et 5.2b).



FIGURE 5.2: Projection des réplicons normés par genre bactérien \bar{V}_{genre}^{R} . Coloration selon le type de réplicon : chromosome (gris clair), plasmide (gris foncé), RECE (bleu).

En prenant un ensemble de cellules D assez large, la SOM parvient à discriminer les plasmides entre eux. Cependant un désavantage non négligeable de la SOM par rapport au MDS est le nombre d'itérations nécessaire (60.000), donc un temps de calcul conséquent, pour produire des résultats robustes.

Concernant les RECE, on peut observer que :

- ► Les RECE se placent généralement au niveau de l'interface entre chromosomes et plasmides et semblent donc posséder des caractéristiques spécifiques les distinguant des plasmides (Figures 5.1c, 5.1d, 5.2a et 5.2b).
- ▶ Quelques RECE montrent une proximité singulière avec les chromosomes.
- ▶ Même si la séparation plasmide/chromosome est nette, il semble exister une proximité entre les plasmides et les chromosomes dont les hôtes appartiennent au même groupe taxonomique, par exemple, pour des groupes tel que les alpha-, bêta- ou gamma-protéobactéries (Figures 5.3a et 5.3b).



FIGURE 5.3: Représentation des lignées bactériennes sur les projections de l'ensemble des réplicons V^R .

Coloration selon la taxonomie de la bactérie-hôte : Acidobactéries (jaune transparent), Actinobactéries (bleu), Bacteroidetes (vert émeraude), Chlamydiae (orange foncé), Chlorobi (rose pâle), Chloroflexi (orange pâle), Cyanobactéries (cyan), Deinococcus-Thermus (bleu transparent), Firmicutes (jaune), Fusobactéries (vert), Planctomycètes (beige clair), Protéobacteries : Alpha (rose), Bêta (jaune pâle), Delta (gris foncé), Epsilon (violet pâle), et Gamma (vert clair), Spirochètes (rouge), Ténéricutes (bleu pâle), Thermotogae (marron foncé), Autres (gris pâle).

Cette proximité entre chromosomes et plasmides d'un même groupe taxonomique témoigne du flux intra-génomique de gènes liés aux STIG entre plasmides et chromosomes.

5.3.2 Graphes

En utilisant l'éq. 3.13, V^R et \bar{V}^R_{genre} sont transformés en graphes bipartites, qui sont ensuite visualisés par des techniques dédiées de visualisation des graphes. Elles consistent à représenter les nœuds et arcs/arêtes des graphes par des points et flèches/traits entre les points dans un espace à deux ou trois dimensions. L'épaisseur des traits peut être une fonction de la pondération. Des algorithmes de topologie ou *layout* organisent la disposition des nœuds dans l'espace selon différentes stratégies. L'une d'entre elles est le *force-directed layout* où un système de forces est appliqué à chaque nœud en fonction de ses arêtes. Le graphe est alors mis en mouvement jusqu'à ce qu'il atteigne un état stable. Ce dernier dépend de la connectivité inter-nœud et peut être assimilé, par analogie, à l'action d'un champ de force (gravité, interaction électrostatique,...) sur une molécule, lui faisant adopter sa configuration la plus stable. L'algorithme que nous avons employé pour visualiser nos graphes est *ForceAtlas2* [Jacomy et al., 2014].

5.3.2.1 Graphes : résultats

L'analyse des graphes formés pour V^R (Figure 5.4) et des sous-graphes des différents groupes taxonomiques contenant des espèces multipartites (Figure 5.5) apporte un autre point de vue que les projections en permettant d'analyser plus précisément les différentes interconnections de chaque réplicon par rapport aux clusters de protéines. L'utilisation du logiciel de visualisation Gephi, par ses nombreuses fonctionnalités, a apporté une contribution non négligeable dans la fouille de données des graphes.

La quasi-majorité des réplicons sont interconnectés, témoignant de l'hérédité partagée des gènes des STIG présents dans les génomes bactériens. Plasmides et chromosomes semblent cependant clairement appartenir à des communautés différentes. Il existe néanmoins pour certains phyla une interconnection forte entre chromosomes et plasmides. Est passé en revue ci-dessous les observations tirées de l'analyse des graphes pour les génomes multipartites. À cause de la grande complexité des données, seulement les observations les plus marquantes visuellement sont rapportées.

- Protéobactéries Différents niveaux de connections des RECE avec les plasmides et chromosomes existent chez les protéobactéries (détaillés ci-dessous). Les RECE des Protéobactéries sont interconnectés via certains clusters remarquables, notamment les clusters annotés IciA, Lrp, FtsE, AcrA ou HN-S (Figure 5.6). De manière générale, les RECE fortement interconnectés avec les plasmides le sont via les clusters annotés ParA/ParB, Rep et en lien avec les systèmes PSK.
 - Alpha Les RECE sont liés pour part à de nombeux clusters plasmidiques dont ceux annotés ParA, ParB et RepC, témoignant de la proximité avec les plasmides de type RepABC présents chez les Alphaprotéobactéries. Cependant, ilsi sont liés à des clusters typiquement chromosomiques, FtsK (pour les RECE de Brucella et d'Agrobacterium) et ParE (pour les RECE de Brucella) notamment. Il existe des clusters spécifiques des RECE, annotés TyrK, FtsE ou AcrA particulièrement. Enfin, l'interconnection des RECE de Paracoccus et de Asticcacaulis avec des réplicons chromosomiques est surprenante (Figure 5.5i). Les RECE des Brucella forment un groupe distinct alors que ceux de Sphingobium semblent plus interconnectés avec des plasmides.
 - Bêta Les RECE se distinguent clairement des chromosomes et des plasmides en formant un groupe homogène distinct (Figure 5.5j). Ils partagent différents clusters avec les plasmides tels que ceux annotés parA ou Xer, entre autres. Les RECE sont interconnectés aux chromosomes via de nombreux clusters tels que FtsI, IciA ou Rob. Il existe des différences entre les RECE I et les RECE II des Burkholderia, ces derniers semblant plus interconnectés aux plasmides.Il est à noter que les RECE I des Burkholderia sont liés à des clusters annotés DnaG connectés aux chromosomes (sauf ceux des Burkholderia), témoignant d'un probable transfert ancestral du chromosome vers le RECE I.
 - **Gamma** Les RECE sont connectés à l'interface des chromosomes et plasmides (Figure 5.5k). Ceux de *Pseudoalteromonas* sont moins interconnectés avec les RECE des *Vibrio/Aliivibrio* et *Photobacterium* soulignant leur origine distincte.



(A) Type de réplicons.

(B) Phylum/Classe des bactéries-hôtes.

FIGURE 5.4: Visualisation des réplicons par graphes bipartites.

5.4a. Selon le type de réplicon : chromosome (gris clair), plasmide (gris foncé), RECE (bleu). 5.4b. Selon la taxonomie de la bactérie-hôte : Acidobactéries (jaune transparent), Actinobactéries (bleu), Bacteroidetes (vert émeraude), Chlamydiae (orange foncé), Chlorobi (rose pâle), Chloroflexi (orange pâle), Cyanobactéries (cyan), Deinococcus-Thermus (bleu transparent), Firmicutes (jaune), Fusobactéries (vert), Planctomycètes (beige clair), Protéobacteries : Alpha (rose), Bêta (jaune pâle), Delta (gris foncé), Epsilon (violet pâle), et Gamma (vert clair), Spirochètes (rouge), Ténéricutes (bleu pâle), Thermotogae (marron foncé), Autres (gris pâle).

- Acidobactéries Le RECE de *Candidatus* Chloroacidobacterium thermophilum est interconnecté aux chromosomes par ParA/ParB, DnaB, AcrA, ScpB principalement, et n'est pas interconnecté avec les plasmides des Acidobactéries. Il est cependant connecté à des plasmides d'autres phylum *via* des clusters Rep, Helicase ou PSK principalement.
- Actinobactéries Le RECE de *Nocardiopsis dassonvillei* est interconnecté aux chromosomes *via* de nombreux clusters annotés principalement FtsE, FtsK, FtsW, FtsI, MinD et ParA.
- **Bacteroidetes** Les RECE de *Prevotella* sont interconnectés aux chromosomes *via* de très nombreux clusters et ont peu d'interconnections avec les plasmides.
- Chloroflexi Les RECE de Sphaerobacter thermophilus et Thermobaculum terrenum sont interconnectés aux chromosomes et aux plasmides via de nombreux clusters. En particulier, le RECE de S. thermophilus est lié aux clusters DnaA (GI : 269929006), DnaG et FtsI.
- **Cyanobactéries** Deux configurations distinctes sont trouvées chez les RECE de Anabaena et Cyanothece : le RECE de Anabaena est assez interconnecté aux chromosomes via des clusters annotés FstI, DnaG, FtsE, CpbA, MreB, ParA entre autre. Inversement le RECE de Cyaothece interconnecté aux plasmides via des clusters de type Xer, ParA et helicase notamment.
- **Deinococcus-Thermus** Le RECE de *Deinococcus* est lié à huit clusters annotés ParA, ParB, FtsE, Helicase, XerC et HupB. Par le nombre de clusters et par son interconnection, Il est beaucoup plus proche des autres plasmides de *Deinococcus*. Il est cependant interconnecté aux chromosomes *via* les clusters FtsE et HupB notamment et est lié à un nombre (trois) relativement élevé de cluster FtsE.
- **Firmicutes** Le RECE de *Butyrivibrio proteoclasticus* est connecté aux plasmides *via* des clusters annotés Rep et PSK principalement et est connecté aux chromosomes *via* ceux annotés DnaB, Rob, FtsE, Xer. pCY186, un plasmide de *B. proteoclasticus* est attaché au même cluster DnaB, ainsi qu'à deux clusters annotés DnaA (GI : 302668636 et 302668625).
- **Spirochètes** Les RECE de *Leptospira* sont liés chacun à trois ou qutre clusters de protéines pour un total de six clusters différents. Les clusters en commun sont annotés ParA et ParB et ceux n'étant pas présents sur tous les RECE sont annotés XerD et HupB. Le cluster annoté ParA est lié exclusivement aux chromosomes des spirochètes, avec une exception : un plasmide de *Turneriella*. Aucun des six clusters n'est lié à un autre plasmide des Spirochètes. Les RECE possèdent de plus deux gènes appartenant à ce cluster.



FIGURE 5.5: Visualisation par graphe bipartite des réplicons $V^{R_{taxon}}$ selon la taxonomie de leur hôte (phylum ou classe pour les Protéobactéries). Coloration selon le type de réplicon : chromosome (gris clair), plasmide (gris foncé), RECE (bleu). Les clusters de protéines (vert foncé) sont visibles pour certaines résolutions.



FIGURE 5.6: Pour centages, par type de réplicon, de connexion aux 40 clusters de proté ines les plus connectés aux observations annotées en tant que RECE de $\bar{V}^{R_{\rm Protéobactéries}}_{genre}.$

Pour chaque cluster (ordonnée), les pourcentages de connexion (abcisse) sont représentés pour les observations annotées en tant que RECE (bleu), chromosome (gris clair) et plasmide (gris foncé).

Dans l'interprétation visuelle des graphes obtenus par ForceAtlas2, un des biais vient du positionnement de certains plasmides faiblement connectés dans certains groupes de chromosomes hautement connectés s'ils sont liés à des clusters de protéines communs. Il est de plus difficile de comparer la pertinence visuelle des graphes bipartites spatialisés avec les visualisations obtenues par projection.

5.3.2.2 Logiciels utilisés

- **Gephi** [Bastian et al., 2009] logiciel "open source" de graphes, a été utilisé pour spatialiser (en utilisant ForceAtlas 2) et visualiser les graphes.
- **Python** a été utilisé pour réaliser les nombreux workflows, préparer les données, préparer les fichiers *graphml* utilisés par Gephi.

5.3.3 Visualisation : discussion

La projection des réplicons selon leurs STIG par MDS et SOM et la visualisation par graphe montrent que les réplicons s'organisent par groupe taxonomique et selon leur type. Même si les chromosomes et plasmides forment des groupes et communautés distinctes, ils partagent tout de même de nombreux gènes des STIG, témoignant de phénomènes d'échange génétique inter-réplicons dans les génomes bactériens. Les RECE semblent de plus posséder une distribution des gènes des STIG les distinguant des chromosomes et plasmides, et se positionnent à l'interface entre chromosomes et plasmides. Il existe différents degrés d'interconnection des RECE avec les chromosomes et les plasmides selon les RECE. Cependant, compte tenu de la grande complexité des données, ces premiers résultats ne permettent pas d'identifier avec certitude des caractéristiques spécifiques des RECE.

5.4 Classification non-supervisée

L'analyse de clustering des réplicons bactériens a pour objectif principal d'étudier le comportement des RECE par rapport aux autres réplicons. Le positionnement de manière stable d'un RECE dans un groupe de chromosomes ou un groupe de plasmides indiquerait une proximité avec les membres de ce groupe. À l'inverse un cluster homogène de RECE peut être le témoin de leur spécificité et les différencier en tant qu'espèce moléculaire distincte des autres réplicons.

5.4.1 Algorithmes de clustering

Différents algorithmes de clustering ont été testés, les critères retenus étant le temps d'exécution, l'implémentation ou l'implémentabilité, les performances mesurées et la stabilité. Pour prendre en compte la très haute dimensionalité des données, différentes approches ont été suivies :

- i) Algorithmes de clustering spécifiques pour des données en grandes dimensions
- ii) Réduction de dimension initiale par projection des données puis clustering dans ce nouvel espace
- iii) Clustering des graphes bipartites de V^R et \bar{V}^R_{qenre} .

Les méthodes de réduction de dimensions testées sont les mêmes que précédemment (Table 5.4) avec q, le nombre de dimensions, variable. Soient Cl un clustering de C clusters, V un ensemble d'observations et d une distance. La Table 5.6 introduit brièvement les principaux algorithmes de clustering utilisés.

TABLE 5.6: Principaux algorithmes de clustering utilisés dans l'analyse des données.

K-Means	[Hartigan, 1975; Hartigan and Wong, 1979] Algorithme de clustering des plus populaires. Le principe original consiste à minimiser l'erreur $E = \sum_{C \in Cl} \sum_{r \in C} d(v_r, \bar{v}^C)$. Les paramètres de l'algorithme étant la distance d , le nombre de clusters k , le nombre d'itérations et d'exécutions de l'algorithme et son mode de convergence. De nombreuses variantes de l'algorithme existent [Gan et al., 2007].
Classifieurs Hiérarchiques Agglomératifs (CHA)	Algorithmes partant d'un état initial où chaque observation est considérée comme une feuille (ou nœud terminal) d'un arbre, puis formant des clusters d'observations en fusionnant de façon itérative les feuilles de cet arbre [Gan et al., 2007]. Différents critères peuvent être utilisés pour choisir la fusion optimale entre deux nœuds, comme par exemple la distance minimale, maximale ou moyenne entre les observations de deux nœuds, la distance entre les centroïdes des nœuds, ou la méthode WARD (<i>cf.</i> ci-dessous).
WARD	[Ward Jr, 1963; Ward Jr and Hook, 1963] Algorithme de la famille des classifieurs hiérarchiques agglomératifs. À partir d'un ensemble de clusters intermédiaires, le principe de cette famille d'algorithmes est de fusionner deux clusters si et seulement si un certain critère objectif est minimal [Gan et al., 2007]. La méthode de Ward consiste à regrouper deux clusters intermédiaires d'un clustering Cl_i de façon à ce que l'inertie intra-cluster $I = \frac{1}{ Cl_i } \sum_{C \in Cl} \sum_{r \in C} d^2(v_r, \bar{v}^C)$ soit minimisée.
DBSCAN	[Ester et al., 1996] Algorithme de clustering par densité. D'une manière itérative, une observation r est ajoutée à un cluster existant C si ceux-ci sont joignables par densité selon d , c'est-à-dire s'il existe une observation $r_i \in C$ telle que $d(v_r, v_{r_i}) \leq v_{seuil}$. Les paramètres de l'algorithme étant d , v_{seuil} et le nombre de points minimum d'un cluster. Cet algorithme s'applique cependant difficilement aux données de dimensionalité élevée car elles sont alors éparpillées dans l'espace de représentation [Gan et al., 2007] (voir aussi §3.4.8).
SUBCLU	[Kailing et al., 2004] Algorithme de clustering des données de dimensionalité élevée. Le principe est d'exécuter une procédure DBSCAN dans les différents sous-espaces où les données sont joignables par densité. Les différents sous-espaces sont sélectionnés de façon similaire à l'algorithme <i>a priori</i> [Agrawal and Srikant, 1994]. Les paramètres de SUBCLU sont les mêmes que ceux de DBSCAN.

INFOMAP

[Rosvall and Bergstrom, 2008] Algorithme de détection de communautés d'un graphe G. Les données étant transformées en graphe bipartite par la relation 3.13. Dans l'algorithme INFOMAP, une communauté est définie de façon assez similaire à celles de l'algorithme MCL : un trajet aléatoire dans un graphe aura plus tendance à rester au sein d'une communauté que d'en sortir, chaque étape du trajet étant décrit par une séquences d'étapes représentant les différents nœuds déjà traversés et par différentes probabilités d'occurence de la prochaine étape. INFOMAP semble de plus être un des algorithmes les plus précis quand à la découverte de communautés au sein d'un graphe [Lancichinetti and Fortunato, 2009], les paramètres de l'algorithme étant le nombre de trajets aléatoires. Il est intéressant de constater qu'en utilisant un graphe bipartite comme structure, les observations ainsi que les variables sont clusterisées.

La stratégie finalement retenue a été : i) d'utiliser une procédure de réduction de dimensions par ACP sur V^R et \bar{V}^R_{genre} puis un clustering des données projetées par l'algorithme de WARD, ii) d'exprimer V^R et \bar{V}^R_{genre} sous forme de graphes bipartites par l'éq. 3.13 puis d'utiliser l'algorithme INFOMAP. Ces deux procédures donnent les meilleurs résultats en terme de performance et présentent les meilleurs compromis en terme de temps d'exécution et d'implémentation.

En effet, l'algorithme du K-Means nécessite de nombreux runs pour présenter des résultats performants et, d'un point de vue pratique, est plus lent que WARD. Par ailleurs, une méthode de classification hiérarchique comme WARD peut mieux correspondre à l'organisation taxonomique sous-jacente des données. Additionnellement, une procédure de CHA est effectuée en utilisant la distance *cosine*, choix motivé par l'obtention de projections pertinentes par MDS en utilisant cette distance. Les résultats obtenus témoignent d'une structuration des plasmides selon la taxonomie mais se révèlent incapable de structurer les chromosomes (Table 5.7). L'approche de SUBCLU, bien qu'intéressante, était infaisable en pratique car elle requiert un nombre considérable de procédures DBSCAN pour les données utilisées et donc un temps de calcul (de l'ordre de la semaine) qu'il n'était pas possible d'envisager. De par la nature des données, une représentation en graphe bipartite semblait être une approche judicieuse : l'interconnection des réplicons via les clusters protéiques est un critère plus intuitif que la mesure d'une distance inter-réplicons. En comparaison avec d'autres algorithmes de détection de communautés (MCL, WalkTrap, Edges Betweenness, Méthode de Newman, Multilevel) [Coscia et al., 2011], INFOMAP semble produire de meilleurs résultats (résultats non montrés). La projection par ACP a été favorisée pour sa rapidité, sa popularité, l'absence de paramètres nécessaires et sa capacité à décorréler les variables Duby and Robin, 2006]. L'utilisation de MDS et de la distance cosine produit, avec WARD, des clusters pour lesquels une structuration des réplicons d'un point de vue taxonomique et selon leurs types est retrouvée, mais avec des scores de stabilité très bas ($\triangle^{Kl} < 0.5$) rendant les clusters non interprétables en pratique. Bien qu'INFOMAP semble présenter des sérieux avantages, la procédure ACP+WARD a été conservée pour deux raisons : i) elle présente des résultats performants en particulier pour la classification de réplicons qui sont liés à de nombreux clusters de protéines. ii) L'organisation des réplicons est fondée sur des critères de distance entre réplicons. Il est alors important de confronter les résultats de ACP+WARD avec ceux obtenus avec INFOMAP.

	\mathbf{Indice}^{a}	ACP+WARD		D INFOMAP		CHA	
Données		V^R	\bar{V}^R_{aenre}	V^R	\bar{V}^R_{aenre}	V^R	\bar{V}^R_{aenre}
$Paramètres^b$		k : 200	k : 100	itérat	ion : 500	k : 200	k : 100
Nombre de clusters ^c		cp:30 175	cp:15 75	222	77	174	67
Variance expliquée (ACP)		0.57%	0.58%	220		114	01
Critère de stabilité ^{d} : \triangle^{Kl}		0.85	0.74	0.82	0.76	0.79	0.70
	homogeneity	0.93	0.83	0.80	0.63	0.90	0.83
Type de réplicons	completeness	0.25	0.20	0.15	0.15	0.18	0.21
	V-measure	0.43	0.32	0.25	0.24	0.31	0.34
	homogeneity	0.93	0.80	0.93	0.69	0.62	0.59
Phylum des chromosomes	completeness	0.35	0.40	0.60	0.61	0.65	0.65
	V-measure	0.51	0.53	0.73	0.65	0.64	0.62
	homogeneity	0.93	0.80	0.85	0.64	0.60	0.55
Classes des chromosomes	completeness	0.16	0.58	0.80	0.82	0.93	0.89
	V-measure	0.66	0.67	0.82	0.72	0.73	0.68
Phylum des plasmides	homogeneity	0.06	0.01	0.88	0.78	0.85	0.68
	completeness	0.16	0.14	0.33	0.35	0.30	0.32
	V-measure	0.08	0.02	0.48	0.48	0.45	0.43
	homogeneity	0.07	0.02	0.84	0.74	0.81	0.68
Classes des plasmides	completeness	0.28	0.36	0.43	0.51	0.41	0.49
	V-measure	0.12	0.03	0.57	0.60	0.54	0.57

TABLE 5.7: Évaluation des procédures de clustering de V^R et \bar{V}^R_{genre} .

 $^a~V\mbox{-}measure$ calculées selon l'éq. 5.12.

 b k, nombre de clusters en *input* et cp, nombre de composantes principales retenues pour la classification

par WARD.

^c Nombre de clusters obtenus par les algorithmes.

 d Critère de stabilité \triangle^{Kl} calculé par l'éq. 5.15.

L'objectif de cette approche est d'estimer l'organisation générale des réplicons selon leurs distributions en protéines similaires, liées aux STIG. D'un point de vue méthodologique, différentes approches complémentaires auraient pu être des choix judicieux :

- Tester les algorithmes de clustering dans des sous-espaces [Gan et al., 2007], de type SUBCLU, comme par exemple CLIQUE, MAFIA, PROCLUS [Gan et al., 2007; Han et al., 2012]
- Tester des approches de discrétisation d'attributs [Witten et al., 2011]
- Faire une étude comparative plus exhaustive sur les méthodes de projection couplées à des algorihmes de clustering classiques
- Tester des approches expérimentales et/ou novatrices de détection de communautés

5.4.1.1 Logiciels utilisés

Scikit-learn librairie de Machine learning de Python utilisée pour le calcul des projections par ACP, MDS, pour les procédures de clustering par les algorithmes DBSCAN, CHA, et WARD, ainsi que pour le calcul des indices de V-measure (homogeneity, completeness).
- **Pycluster** [de Hoon et al., 2004] utilisé pour les procédures de clustering par K-Means, notamment.
- **igraph** [Csardi and Nepusz, 2006] a été utilisé *via* Python pour l'utilisation des algorithmes de détection de communautés : INFOMAP notamment.
- **Python** utilisé pour réaliser les workflows, préparer les données, calculer les distances inter-observations, ainsi que pour réaliser une implémentation de l'algorithme SUBCLU.
- C++ a été utilisé en parallèle avec Python pour coder certaines parties de divers algorithmes.

5.4.2 Résultats du clustering

Les résultats de l'étude de stabilité et de la séparation des différents ensembles étudiés sont présentés Table 5.7. L'organisation et les spécificités des différents clusters contenant des RECE sont de plus présentées Table 5.8. Pour la procédure ACP+WARD, différentes valeurs de k et de nombre de composantes principales cp ont été testées. cp a été choisi en fonction de l'évolution de la variance expliquée par les composantes et de la stabilité des clusters obtenus et k a été choisi en fonction des indices de stabilité. De même, le choix de k pour la procédure CHA s'est effectué *via* les scores de stabilité.

Les trois méthodes produisent des clusters stables et exploitables dans leur ensemble. Ces procédures parviennent à séparer clairement les plasmides des chromosomes, comme attendu. Cependant, seul ACP+WARD et INFOMAP arrivent à séparer de façon efficace les chromosomes selon leur phylum ou classe taxonomique, ce qui indique une cohérence biologique des résultats. L'échec de CHA à structurer les chromosomes est probablement dû au fait qu'il existe des différences trop importantes de densité entre les groupes de chromosomes et de plasmides et qu'ainsi, les chromosomes sont considérés comme un groupe homogène par rapport aux clusters de plasmides et sont donc difficilement divisés par l'algorithme. Il est intéressant de constater qu'INFOMAP arrive à identifier des clusters ayant une correspondance assez forte avec des groupes taxonomiques (indice de *completeness* élevé). L'ensemble des clusters de réplicons obtenus par $f_{INFOMAP}(V^R)$ sont présentés en Annexe D.

L'utilisation de données normées selon le genre taxonomique et le type des réplicons \bar{V}_{genre}^R tend à produire des clusters moins stables et à faire baisser les valeurs des indices d'homogeneity. Cependant les résultats pour V^R et \bar{V}_{genre}^R sont globalement similaires.

La procédure ACP+WARD a tendance à grouper les plasmides au sein de clusters larges et stables, ce qui s'explique par le fait que la distance calculée $d(v_{r_1}, v_{r_2})$ entre deux plasmides r_1 et r_2 par WARD est davantage sensible au phénomène du *fléau de la dimensionalité* (§3.4.8) que dans le cas des chromosomes en général. À l'inverse, INFOMAP et CHA sont capables de détecter des groupes de plasmides fortement liés à un faible nombre de clusters de protéines communs et corrélés à la taxonomie (Table 5.7). Un des bruits rencontrés dans les clusters d'INFOMAP est qu'un plasmide ou groupe de plasmides isolé(s) peut être intégré à l'intérieur de communautés de chromosomes s'ils sont fortement reliés à des clusters de protéines caractéristiques de ces communautés (ce qui explique les meilleurs scores de ACP+WARD pour la séparation des réplicons). Tout dépend alors de ce qui est considéré comme distance pertinente dans la comparaison des réplicons.

Le principal résultat de cette analyse est cependant la structuration des RECE dans les clusters (Table 5.8). Dans l'ensemble, les RECE se distinguent nettement des autres réplicons en présentant des caractéristiques singulières, ce qui produit des clusters spécifiques de RECE. Ceci est très net chez Vibrio, Leptospira, Brucella ou pour les Burkhoderiales (avec CHA) par exemple. Les RECE de Vibrio/Aliivibrio sont présents dans un unique cluster. Enfin, ceux de Sinorhizobium sont associés avec des plasmides d'espèces de Rhizobiales. Pour les RECE des Burkholderiales et d'Agrobacterium, les clusters obtenus contenant ces RECE sont instables et les scores individuels \triangle^e sont faibles, rendant les résultats peu interprétables. Des clusters des RECE des Burkholderiales stables et homogènes peuvent être cependant obtenus en utilisant uniquement $V^{R_{\text{Betaproteobactéries}}}$, témoignant de la spécificité de ces RECE. Certains mégaplasmides des Burkholderiales appartiennent cependant à ces mêmes clusters. Les RECE II des Burkholderiales, plus interconnectés aux plasmides, ont plus tendance à être dans des clusters avec d'autres plasmides. Ces éléments sont en faveur d'une transition mégaplasmide \Rightarrow RECE de type RECE II \Rightarrow RECE I chez les Burkholderiales.

INFOMAP tend à grouper une majorité de RECE dans des communautés de chromosomes, indiquant que chromosomes et RECE possèdent des caractéristiques communes. À l'inverse, ACP+WARD et CHA ont plus tendance à placer les RECE dans des clusters de plasmides, témoignant d'une différence significative, dans l'ensemble, entre le nombre de clusters de protéines liés aux RECE et aux chromosomes. Cependant on constate que certains RECE sont plus proches des chromosomes alors que d'autres sont liés de manière forte à des plasmides. Pour les trois méthodes, les RECE de Anabaena, Asticaccaulis, Paracoccus et Prevotella sont groupés de manière stable avec les chromosomes (Table 5.8). Cette structuration indique clairement que ces RECE possèdent des caractéristiques propres les distinguant des plasmides. À l'inverse, certains RECE, i.e., chez Deinococcus, Rhodobacter, Leptospira et Butyrivibrio, sont classés de manière stable avec les plasmides indiquant le partage de caractéristiques communes. Enfin, un autre résultat intéressant est l'organisation (inattendue) par IN-FOMAP et par CHA des plasmides selon leur groupe taxonomique révélant un lien entre organisation des STIG plasmidiques et taxonomie des bactéries les hébergeant. Pour un niveau taxonomique donné (phylum ou classe), les plasmides sont organisés en différents groupes (indice de *completeness* faible), indiquant que ces réplicons peuvent être caractérisés par une structuration taxonomique et fonctionnelle.

Les visualisations et les clusterings de V^R et V^R_{genre} montrent que la structuration des réplicons bactériens en fonction de leurs STIG dépend de la taxonomie et du type des réplicons. Il est de plus clair qu'une partie des RECE se distingue, par leurs STIG, nettement des chromosomes et des plasmides. L'ensemble des RECE semble de plus posséder un ou plusieurs gènes liés aux STIG inhabituels pour les plasmides.

C : nombre de clusters regroupant les RECE d'un genre bactérien donné. BHIw : valeur de l'indice BHIw (éq. 4.13) concernant le phylum de l'ensemble des réplicons de ces clusters. %chr, %pl et %RECE : pourcentage de représentation du type des réplicons présents dans ces clusters. $E(\triangle^r)$: valeur moyenne de l'estimateur \triangle^r (relation 5.16) des différents RECE. $\bar{E}(\triangle^C)$: valeur moyenne de l'estimateur \triangle^C (relation 5.14) des différents clusters, pondérée par la taille des clusters (similaire à la relation 5.15). Surlignage : genres taxonomiques où les réplicons apparaissent clairement proches des chromosomes (orange) ou des plasmides (magenta) pour les deux procédures. Le nombre d'étoiles est un indice subjectif de la confiance accordée à ces résultats.

(A) INFOMAP						(b) PCA+WARD									
Genre	\mathbf{C}	BHIw	%chr	%pl	%RECE	$E(\triangle^r)$	$\bar{E}(\triangle^C)$	Genre	С	BHIw	%chr	%pl	%RECE	$E(\triangle^r)$	$\bar{E}(\Delta^C)$
A grobacterium	3	0.9	67	16	17	0.45	0.55	Agrobacterium	2	0.94	0	71	29	0.85	0.58
Aliivibrio	1	1.0	0	0	100	0.97	0.94	Aliivibrio	1	1.0	0	61	39	0.95	0.66
Anabaena **	1	1.0	98	0	2	1.0	0.99	Anabaena	1	0.97	98	0	2	0.0	0.84
Asticcacaulis **	1	1.0	96	3	1	0.8	0.96	Asticcacaulis **	1	1.0	88	4	8	1.0	0.88
Brucella	1	1.0	0.0	5	95	1.0	0.84	Brucella	2	0.96	0	57	43	0.93	0.53
Burkholderia	2	0.99	63	2	17	0.46	0.61	Burkholderia	7	0.97	0	29	71	0.8	0.68
Butyrivibrio *	1	1.0	0	50	50	1.0	0.87	Butyrivibrio *	1	0.27	0	99	1	0.96	0.98
Chlora cidobacterium	1	0.82	92	8	1	1.0	0.95	Chlora cido bacterium	1	0.27	0	99	1	0.96	0.98
Cupriavidus	1	0.99	71	11	18	0.45	0.59	Cupriavidus	1	1.0	0	25	75	0.89	0.63
$Cyanothece \ *$	1	0.89	0	94	6	0.5	0.64	$Cyanothece \ *$	1	0.27	0	99	1	0.96	0.98
Deinococcus **	1	0.71	0	96	4	1.0	0.76	$Deinococcus \ ^*$	1	0.27	0	99	1	0.96	0.98
Ilyobacter	1	0.82	0.92	8	1	1.0	0.95	Ilyobacter	1	0.27	0	99	1	0.96	0.98
Leptospira *	1	1.0	0	13	88	1.0	0.94	Leptospira *	1	27	0	99	1	0.96	0.98
No cardiops is	1	0.97	90	9	0	1.0	0.96	No cardiops is	1	0.58	0	98	2	0.67	0.4
Ochrobactrum	1	1.0	0	5	95	1.0	0.84	Ochrobactrum	1	1.0	0	0	100	1.0	1.0
Paracoccus **	1	1.0	96	3	1	0.8	0.96	Paracoccus **	1	1.0	88	4	8	1.0	0.88
Photobacterium	1	0.99	0.71	11	18	0.45	0.59	Photobacterium	1	1.0	0	0	100	0.51	0.54
Prevotella ***	1	0.95	95	2	2	1.0	0.96	Prevotella **	2	1.0	95	0	5	0.5	0.71
Pseudoal teromonas	1	0.99	97	3	1	0.98	0.82	Pseudoalteromonas	2	0.28	0	98	1	0.96	0.98
Ralstonia	1	0.99	71	11	18	0.45	0.59	Ralstonia	3	1.0	0	23	77	0.87	0.77
Rhodobacter *	1	1.0	0	60	40	1.0	0.9	Rhodobacter *	2	0.65	0	94	6	0.72	0.43
Sinorhizobium **	2	0.96	0	98	2	0.67	0.65	Sinorhizobium **	2	0.94	0	79	21	0.64	0.51
Sphaerobacter	1	1.0	0.0	50	50	1.0	0.74	Sphaerobacter	1	0.93	0	8	20	0.75	0.52
Sphingobium	2	0.95	78	20	1	0.65	0.93	Sphingobium	1	1.0	0	61	39	0.95	0.66
Thermobaculum	1	0.82	92	8	1	1.0	0.95	Thermobaculum	1	0.27	0	99	1	0.96	0.98
Variovorax	1	0.99	71	11	18	0.45	0.59	Variovorax	1	1.0	0	33	67	0.5	0.48
Vibrio	1	1.0	0	0	100	0.97	0.94	Vibrio	2	1.0	0	44	56	0.74	0.63

Chapitre 6

Analyse fonctionnelle des réplicons

Pour caractériser les bases biologiques de la différentiation des types de réplicon, nous avons procédé à des analyses complémentaires en ne considérant plus comme variables les clusters d'homologues de protéines mais leurs annotations fonctionnelles. Chaque réplicon est ainsi décrit par une distribution de fonctions des STIG, qui sert de base à la comparaison/discrimination des réplicons. Un avantage de cette modalité de représentation est de réduire la dimensionalité des données. L'inconvénient, cependant, est que l'information liée aux homologies de séquence des protéines est perdue. Nous avons effectué trois types d'analyses fonctionnelles : i) la visualisation des réplicons et des génomes par les fonctions des clusters, ii) des analyses de régression de la distribution des fonctions selon le type de réplicon et de génome, et iii) des analyses de classification des réplicons et génomes.

6.1 Jeux de données et notation

6.1.1 Notations

Soient une protéine p et son annotation Ann(p) définie par la relation 4.10, et un groupe de protéines C. On définit Ann(C), l'annotation de C par :

$$Ann(C) = Ann(p) \iff N^C_{Ann(p)} = max\{N^C_{Ann(p_i)} | p_i \in C\}$$
(6.1)

où $N_{Ann(p)}^C$ est le nombre de fois que Ann(p) apparaît dans C, tel que :

$$N_{Ann(p)}^{C} = |\{Ann(p_i)| \ p_i \in C \text{ et } Ann(p) = Ann(p_i)\}|$$

$$(6.2)$$

Soit un cluster de protéines Cl. On définit par F, son ensemble fonctionnel, tel que :

$$F^{Cl} = \{Ann(C) \mid C \in Cl \text{ et } Ann(C) \text{ unique}\}$$

$$(6.3)$$

Soient r un réplicon et v_r son vecteur associé pour un clustering de protéines Cl donné, défini par l'éq. 5.4. On définit alors v_r^f , son vecteur fonctionnel tel que :

$$v_r^f = (N_{f_1}^r, \dots, N_{f_{|F|}}^r), \ f \in F^{Cl}$$
(6.4)

où N_f^r est défini par :

$$N_f^r = \sum_{\substack{C \in Cl\\Ann(C) = f}} N_C^r \tag{6.5}$$

avec N_C^r défini par l'éq. 5.4.

Pour un génome $g = \{r_1, ..., r_{|g|}\}$ (relation 5.2), on définit v_f^g par :

$$v_f^g = (N_{f_1}^g, ..., N_{f_{|F|}}^g), \ f \in F$$
(6.6)

où N_f^g est défini par :

$$N_f^g = \sum_{r \in g} N_f^r \tag{6.7}$$

Pour des ensembles de réplicons R et de génomes G, on définit $V_f^R = \{v_f^r | r \in R\}$ et $V_f^G = \{v_f^g | g \in G\}$ similairement à l'éq. 5.5, ainsi que \bar{V}_{f,n_tax}^R similairement à l'éq. 5.9. \bar{V}_{f,n_tax}^G est de plus défini par :

$$\bar{V}_{f,n_tax}^G = \bar{V}^{Kl_{n_tax}^{G^{\{monopartite\}}}} \cup \bar{V}^{Kl_{n_tax}^{G^{\{multipartite\}}}}$$
(6.8)

où $G^{\{monopartite\}}$ est l'ensemble des génomes ne contenant pas de RECE et $G^{\{multipartite\}}$ est l'ensemble des génomes contenant au moins un réplicon annoté comme RECE.

Les annotations utilisées pour les protéines de P_{ref} correspondent aux annotations de KEGG (Annexe B) et à nos annotations des groupes ACLAME (Annexe C).

6.1.2 Dimension des données

Les 6096 clusters de protéines (Cl) obtenus par TRIBE-MCL avec une granularité gr = 4 ont été annotés par **117** fonctions (71 de KEGG et 46 de ACLAME). **2720** génomes bactériens ont de plus été formé à partir des données. Similairement aux analyses précédentes, l'indice taxonomique n_tax de normalisation des données choisi est le genre. Les caractéristiques des données sont présentées Table 6.1.

Ensemble de données	Matrice	Taille
Cl	-	6096
F^{Cl}	-	117
V_f^R	$M^{V_f^R}$	(4928, 117)
$ar{V}^R_{f,genre}$	$M^{\bar{V}^R_{f,genre}}$	(851, 117)
V_f^G	$M^{V_f^G}$	(2720, 117)
$ar{V}^G_{f,genre}$	$M^{\bar{V}_{f,genre}^G}$	(584, 117)

TABLE 6.1: Dimension des données.

6.2 Discrimination fonctionnelle des réplicons

Une ségrégation fonctionnelle d'un certain groupe de réplicons est potentiellement le témoin de la spécificité fonctionnelle de ce groupe par rapport aux autres réplicons. L'objectif sous-jacent est alors d'identifier des groupes de RECE, **témoins de l'adap-tation fonctionnelle de ces éléments**. La discrimination des réplicons selon leur type à partir de V_f^G et V_f^R a donc été recherchée.

Les composantes principales des ACP sur V_f^R et $\bar{V}_{f,genre}^R$ permettent d'expliquer la majorité de la variance des données (Figure 6.1).



FIGURE 6.1: Variance expliquée par les quatre composantes principales des ACP sur V_f^R (rouge), $\bar{V}_{f,genre}^R$ (bleu), $V_{f,genre}^G$ (vert) et $\bar{V}_{f,genre}^G$ (rose).

Les projections des réplicons sur les deux premières composantes sont alors utilisées pour visualiser les données. Une procédure de clustering par WARD est conduite sur les quatre premières composantes des projections de V_f^R et $\bar{V}_{f,genre}^R$. Le choix de k, nombre

de clusters *input* de l'algorithme, est effectué en fonction du critère de stabilité \triangle^{Kl} (éq. 5.15) pour $k \in \{3, 10, 20, 50, 70, 100, 150\}$. Les clusters obtenus sont globalement stables ($\triangle^{Kl} \approx 0, 75$) et sont donc interprétables (Table 6.2).

	\mathbf{Indice}^{a}	ACP+	-WARD
Données		V_f^R	$ar{V}^R_{f,genre}$
Paramètres de WARD ^b		k:50 cp:4	k:20 $cp:4$
Nombre de clusters (ACP)		49	19
Variance expliquée (ACP)		87%	85%
Critère de stabilité $^{c} riangle^{Kl}$		0.80	0.71
	homogeneity	0.85	0.68
Type de réplicon	completeness	0.30	0.23
	V-measure	0.44	0.35
	homogeneity	0.50	0.44
Phylum des chromosomes	completeness	0.27	0.33
	V-measure	0.35	0.38
	homogeneity	0.47	0.37
Classe des chromosomes	completeness	0.36	0.41
	V-measure	0.41	0.39
	homogeneity	0.02	0.02
Phylum des plasmides	completeness	0.10	0.3
	V-measure	0.03	0.03
	homogeneity	0.03	0.02
Classe des plasmides	completeness	0.25	0.28
	V-measure	0.05	0.03

TABLE 6.2: Évaluation des procédures de clustering de V_f^R et $\bar{V}_{f,genre}^R$.

 $^a\,$ V-measure calculée selon l'éq. 5.12.

 b k, nombre de clusters en *input* et cp nombre de composantes principales retenues pour la

classification par WARD.

 c Critère de stabilité \bigtriangleup^{Kl} calculé par l'éq. 5.15.

• Chromosomes, plasmides et RECE sont discriminés fonctionnellement selon les annotations de leurs protéines des STIG. Les chromosomes et plasmides forment deux ensembles distincts (Table 6.2; Figures 6.2 et 6.3), confirmant les spécificités fonctionnelles respectives de leurs STIG.



FIGURE 6.2: Projection des réplicons selon les quatre premières composantes de l'ACP sur V_f^R ou V_f^R . Chromosome : gris clair. Plasmide : gris foncé. RECE : bleu (V_f^R) ou rouge (\bar{V}_f^R) .

• Les plasmides s'organisent en un groupe compact alors que différents sous-ensembles de chromosomes peuvent être identifiés (Figure 6.3). Le regroupement des plasmides peut s'expliquer comme précédemment (Chapitre 5), où les distances interplasmides sont faibles à cause du petit nombre de fonctions présentes chez eux.



FIGURE 6.3: Visualisation des lignées bactériennes sur la projection des réplicons selon les deux composantes principales d'une ACP sur V_f^R (6.3a et 6.3b) et, après normalisation taxinomique, sur \bar{V}_f^R (6.3c et 6.3d).

6.3b et 6.3d : zooms (rectangle) des figures 6.3a et 6.3c, respectivement. Chromosome : gris clair. Plasmide : gris foncé. RECE : bleu (V_f^R) ou rouge (\bar{V}_f^R) .

- Le nombre de variables positives dans un vecteur de réplicon n'est pas le seul facteur expliquant la discrimination des réplicons. En particulier, des RECE liés à peu de fonctions (tels que ceux d'Anabaena, Butyrivibrio ou Ilyobacter) sont très nettement séparés des plasmides présentant un nombre de fonctions similaires (Figure 6.3d). Plus généralement, les RECE se différencient des plasmides et se placent de façon caractéristique par rapport aux chromosomes (Figure 6.2), à l'exception des RECE des Spirochètes (Leptospira; Figure 6.3d). Les RECE possèdent ainsi des spécificités fonctionnelles des STIG caractéristiques, les différenciant d'une part des chromosomes et d'autre part des plasmides.
- Les génomes de *Prevotella* présentent un cas singulier où les chromosomes des espèces à génome multipartite (*P. intermedia* et *P. melaninogenica*) se distinguent de ceux des espèces à génome monopartite (*P. ruminicola* et *P. denticola*) en se rapprochant du groupement compact des plasmides (Figure 6.3b).
- Sur la base de leur distribution de protéines des STIG, plusieurs types de RECE peuvent être caractérisés. Différents niveaux de spécificité semblent exister chez les RECE, les rapprochant soit des chromosomes, soit des plasmides (Table 6.2 et Figure 6.3). Par exemple, les RECE des Alpha-protéobactéries semblent plus proches des chromosomes que ceux de la cyanobactérie *Cyanothece* et de *Candidatus* Chloroacidobacterium thermophilum (Chlorobi) qui se localisent à proximité des plasmides. Plus généralement, les résultats de ces analyses sont globalement cohérents avec les conclusions de l'étude de la totalité des réplicons (Table 5.8). Les RECE "proches" des chromosomes se retrouvent dans des clusters de chromosomes et, de même, les RECE avoisinant des plasmides (celui de *Butyrivibrio* excepté) sont retrouvés dans les clusters de plasmides (Table 6.3).
- Le mode de vie des organismes bactériens influe sur les STIG de leurs réplicons génomiques. Outre les spécificités observables de certains groupes de chromosomes de Protéobactéries, Firmicutes, Actinobactéries...(Figure 6.3a), certains chromosomes présentent un biais de positionnement par rapport à la majorité des chromosomes et se placent à proximité des plasmides en se regroupant à leur voisinage ou avec ces derniers dans des clusters communs. Ces chromosomes appartiennent à des lignées diverses d'espèces endosymbiotiques ou de parasites/pathogènes d'eucaryotes (Rickettsia, Wolbachia, Helicobacter, Chlamydia, Mycoplasma, Borrelia, ...). La plupart ont des tailles relativement réduites $(\simeq 1Mb: Chlamydia, Borrelia...)$, voire très réduites (< 0.5Mb: Candidatus Carsonella, Candidatus Tremblaya...). Les génomes réduits, en adoptant une taille restreinte afin d'optimiser leur développement au sein de l'organisme hébergeant la bactérie et en relation avec la dynamique de population de la bactérie-hôte (diminution de N_e), développent des STIG spécifiques. La taille n'est toutefois pas le seul facteur expliquant cette proximité. Par exemple, les génomes d'*Helicobacter* (epsilon-protéobactéries) possèdent une taille relativement importante dans le contexte des génomes réduits ($\simeq 1.6Mb$) mais sont fortement discriminés en comparaison à des génomes de taille similaire tels que Taylorella (gamma-protéobactéries) ou Prochlorococcus (Cyanobactéries).

Genre	\mathbf{C}	BHIw	%chr	%pl	%RECE	$E(\triangle^r)$	$\bar{E}(\triangle^C)$
A grobacterium	3	0.86	65	15	21	0.95	0.60
Aliivibrio	1	1.00	0	30	70	0.80	0.70
Anabaena	1	0.21	77	19	4	0.68	0.64
A stic ca caulis	1	0.51	99	0	1	1.00	0.60
Brucella	1	0.75	43	25	32	0.98	0.80
Burkholderia	6	0.92	31	27	42	0.85	0.68
Butyrivibrio	1	0.21	77	19	04	0.68	0.64
Chloroacido bacterium	1	0.29	1	99	0	0.90	0.98
Cupriavidus	1	1.00	5	0	95	0.71	0.66
Cyanothece	1	0.29	1	99	0	0.90	0.98
Deinococcus	1	0.29	1	99	0	0.90	0.98
Ilyobacter	1	0.21	77	19	4	0.68	0.64
Leptospira	1	0.29	1	99	0	0.90	0.98
No car diops is	1	0.21	77	19	4	0.68	0.64
Ochrobactrum	1	1.0	90	3	7	0.75	0.39
Paracoccus	1	0.89	92	3	5	0.53	0.32
Photobacterium	1	1.00	0	30	70	0.80	0.70
Prevotella	2	0.34	87	11	3	0.74	0.62
Pseudoalteromonas	1	0.21	77	19	4	0.68	0.64
Ralstonia	2	1.00	11	22	77	0.77	0.69
Rhodobacter	2	0.84	27	41	32	0.97	0.73
Sin or hizobium	2	0.86	25	27	48	0.87	0.76
Sphaerobacter	1	0.35	100	0	0	1.00	0.60
Sphingobium	2	34	64	27	9	0.84	0.64
Thermobaculum	1	0.21	77	19	4	0.68	0.64
Variovorax	1	1.00	0	30	70	0.80	0.70
Vibrio	4	0.97	31	32	37	0.83	0.57

TABLE 6.3: Classification non-supervisée des RECE par ACP+WARD sur V_f^R . Indices identiques à ceux de la Table 5.8. Paramètres de l'analyse ACP+WARD décrits Table 6.2.

6.3 Discrimination fonctionnelle des génomes

La même analyse a été appliquée aux génomes $(V_f^G \text{ et } \bar{V}_{f,genre}^G)$. L'ACP a été utilisée pour projeter les données dans un espace de dimension réduite afin de les visualiser et les clusteriser. Les quatre principales composantes ont de plus été utilisées pour la classification par WARD. Le choix de k pour l'analyse de clustering est fondé sur le critère de stabilité Δ^{Kl} (éq. 5.15) pour $k \in \{3, 10, 20, 50, 70, 100, 150\}$.

	$\mathbf{Indices}^{a}$	ACP+	WARD
Données		V_f^G	$ar{V}^G_{f,genre}$
Paramètres de WARD b		$k:150\\cp:4$	$k:70\\cp:4$
Nombre de clusters ACP		142	60
Variance expliquée ACP		82%	81%
$\mathbf{Crit} \mathbf{\dot{e}re} \ \mathbf{de} \ \mathbf{stabilit} \mathbf{\acute{e}^{c}} \ \bigtriangleup^{Kl}$		0.68	0.58
	homogeneity	0.60	0.39
Type de génome	completeness	0.03	0.02
	V-measure	0.05	0.04
	homogeneity	0.51	0.55
Phylum des génomes	completeness	0.22	0.29
	V-measure	0.31	0.38
	homogeneity	0.51	0.51
Classe des génomes	completeness	0.32	0.39
	V-measure	0.39	0.44

TABLE 6.4: Évaluation du clustering fonctionnel des génomes.

 $^a\,$ V-measure calculée selon l'éq. 5.12.

b k, nombre de clusters en *input* et cp, nombre de composantes principales

retenues pour la classification par WARD.

 c Critère de stabilité \bigtriangleup^{Kl} calculé par l'éq. 5.15.

Les clusters de génomes obtenus par ACP+WARD sont plus instables que ceux obtenus pour les réplicons (Table 6.4) et sont donc difficilement interprétables et potentiellement fortement bruités. Les critères d'homogeneity indiquent de plus qu'il n'y a *a priori* pas de séparation nette détectable entre génomes mono- et multipartites. Cependant, les projections des données selon les quatre composantes principales de l'ACP sont pertinentes compte-tenu de la variabilité expliquée (> 80%, Table 6.4). De plus, les projections de $\bar{V}_{f, denre}^{G}$ montrent des tendances similaires à celles de V_{f}^{G} (Figure 6.4).

- Aucune tendance générale ne caractérise les génomes multipartites dans leur globalité. Les variables utilisées ne permettent pas de mettre en avant une discrimination significative des génomes multipartites par rapport aux génomes monopartites (Table 6.4). D'éventuelles spécificités des génomes qui, à l'échelle des réplicons étaient clairement visibles, sont peut-être "noyées" dans la masse des données génomiques. Alternativement, il est possible que l'on ne dispose de pas suffisamment de variables pour qu'émerge une tendance caractéristique des génomes multipartites.
- Certains génomes, dont une part importante de génomes multipartites, sont cependant clairement discriminés. Les projections indiquent que certains génomes s'écartent de l'ensemble des génomes (Figure 6.4). Ils englobent les génomes multipartites des Bêtaprotéobactéries ainsi qu'une part importante des



FIGURE 6.4: Projection fonctionnelle des génomes selon les deux composantes principales d'une ACP sur V_f^G (A,B) ou $\bar{V}_{f,genre}^G$ (C).

Génomes monopartites : gris foncé ; génomes multipartites : magenta. Bêtaprotéobactéries : kaki, Alphaprotéobactéries : vert. A : Visualisation de la tendance de répartition des Bêta-protéobactéries et de groupes majoritairement formés de génomes monopartites (jaune). B : Visualisation des localisations des génomes monopartites des Bêtaprotéobactéries et des Alphaprotéobactéries (Rhizobiales indiquées par des marqueurs carré). Rectangle rouge à droite de la figure : génomes uniquement composés de plasmides. C : Visualisation des génomes multipartites les plus différenciés parmi les Bêtaprotéobactéries (kaki) et les Alphaprotéobactéries (vert) selon une projection de $\bar{V}_{f,genre}^G$ sur les deux composantes principales.

génomes des Rhizobiales (Alphaprotéobactéries). Ces génomes appartiennent majoritairement à des espèces associées à des plantes, parmi lesquelles des pathogènes mais surtout des mutualistes qui sont connus pour échanger des gènes impliqués dans le développement de l'association symbiotique aux plantes. Au-delà des gènes impliqués dans la symbiose avec un hôte eucaryote, l'écologie particulière de ces organismes (symbiotes des plantes) est corrélée à une adaptation spécifique des STIG de ces génomes. Cette corrélation n'étant pas détectable sur les analyses des réplicons (Figure 6.3), on peut proposer l'hypothèse corollaire que cette adaptation des STIG passe par l'adaptation des réplicons extra-chromosomiques. Ces hypothèses sont toutefois à considérer avec prudence, compte tenu de l'instabilité des données, et seront à confirmer ou infirmer par des analyses complémentaires.

6.4 Analyses par régression et test d'hypothèses

L'identification des différents biais fonctionnels des groupements de réplicons ou de génomes conduit naturellement à tenter d'identifier les fonctions impliquées dans la discrimination de ces groupes. Pour comparer les groupes d'éléments génomiques, nous avons réalisé une régression logistique sur les distributions des 117 fonctions de notre analyse. Deux critères sont considérés : la P_{value} mesurant la capacité d'une fonction à être spécifique d'une classe donnée et l'**Odd-Ratio** évaluant, pour une observation donnée, l'influence de la présence d'une fonction sur l'appartenance de cette observation à une classe donnée.

6.4.1 Régression logistique : principe

Pour un ensemble E d'observations expliquées par n attributs, organisé en deux sousensembles tels que $E = E_{True} \cup E_{False}$, $E_{training} = \{E_{True}, E_{False}\}$ représente une procédure de régression logistique simple, assimilable à une procédure de classification supervisée binaire $f_{reg}^{E_{training}}$ où :

$$f_{reg}^{E_{training}}(o) = \frac{e^{\beta_0 + \sum_{1 \le i \le |v_o|} \beta_i v_o[i]}}{1 + e^{\beta_0 + \sum_{1 \le i \le |v_o|} \beta_i v_o[i]}}, \ o \in E$$
(6.9)

avec v_o , le vecteur associé à l'observation o et β_i , les coefficients de la fonction *logit* définie par :

$$g(o) = ln\left(\frac{f_{reg}^{E_{training}}(o)}{1 - f_{reg}^{E_{training}}(o)}\right) = \beta_0 + \sum_{1 \le i \le |v_o|} \beta_i . v_o[i]$$
(6.10)

Les β_i sont alors estimés afin de proposer le modèle expliquant au mieux les observations de $E_{training}$, de telle sorte que $f_{reg}^{E_{training}}(o) \approx 1$ pour $o \in E_{True}$ et $f_{reg}^{E_{training}}(o) \approx 0$ pour $o \in E_{False}$ [Hosmer Jr et al., 2013]. $f_{reg}^{E_{training}}(o)$ étant inclus dans l'intervalle]0, 1[, cette valeur peut être considérée comme la probabilité qu'o appartiennent à E_{True} [Larose, 2006]. La signifiance du modèle, ou des coefficients β_i , peut alors être estimée via des tests statistiques sur la *deviance* du modèle ou en applicant le test de Wald individuellement sur chaque coefficient [Larose, 2006]. Le test de Wald, en particulier, évalue si un coefficient β_i est significativement différent de 0. Dans de cas d'observations ayant un unique attribut dichotomique (deux valeurs possibles), le rapport de chance (ou **O**dd **R**atio (OR)) [Larose, 2006] est défini par :

$$OR = \frac{\frac{f_{reg}^{E_{training}}(1)}{1 - f_{reg}^{E_{training}}(0)}}{\frac{f_{reg}^{E_{training}}(0)}{1 - f_{reg}^{E_{training}}(0)}} = e^{\beta_1}$$
(6.11)

Un OR de 5,0 pour un modèle de *logit* défini sur un ensemble d'observations univariées prenant leur valeur dans $\{0, 1\}$ indiquera que, selon le modèle de régression, les observations prenant la valeur 1 auront 5 fois plus de chances d'être dans la catégorie E_{True} que les observations ayant comme valeur 0. Pour des observations univariées continues suivant l'hypothèse que β_1 est constant, le rapport de chance e^{β_1} est alors l'augmentation de chance d'appartenir à E_{True} suivant l'augmentation des valeurs des observations d'une unité [Hosmer Jr et al., 2013].

6.4.2 Jeux de données

Différentes classes de référence K_i d'ensembles de réplicons ont été formées (Table 6.5). Les génomes incomplets, ne contenant pas de chromosomes, ne sont pas pris en considération. D'une manière générale, les données sont systématiquement normées par genre taxonomique afin d'éviter les corrélations dues à la sur-représentation de certains genres bactériens. Bien qu'il semble exister plusieurs types différenciés, les RECE ont été rassemblés dans une unique classe de référence K_{RECE} , le nombre d'observations n'étant pas suffisant pour permettre la création de plusieurs classes. La classe de référence $K_{plasmide}$ rassemble l'ensemble des plasmides et la classe K_{chr} regroupe l'ensemble des chromosomes. Les classes $K_{monopartite}$ et $K_{multipartite}$ regroupent, respectivement, l'ensemble des génomes monopartites et celui des génomes multipartites. Soit $G_{alphaproteo}$ et $G_{betaproteo}$ les ensembles des génomes des classes Alphaprotéobactéries et Bêtaprotéobactéries, respectivement. Les classes $K_{alphamono}$, $K_{alphamulti}$, $K_{betamono}$, $K_{betamulti}$ sont alors formées <u>sans</u> normer les observations par genre taxonomique, les espèces étant alors l'objet de la comparaison.

Nous avons donc étudié par régression logistique la pertinence des fonctions des STIG pour représenter ces différentes classes de réplicons. Les chromosomes sont comparés aux plasmides puis aux RECE. Les RECE sont ensuite comparés aux plasmides. Enfin, pour approfondir les résultats de discrimination fonctionnelle des réplicons par ACP+WARD (§6.2), les génomes monopartites et multipartites sont comparés à l'intérieur des Bêtaprotéobactéries et des Alphaprotéobactéries, respectivement.

Classe	$\mathbf{Donn\acute{e}es}^a$	\mathbf{Taille}^{b}
K_{RECE}	$ar{V}^{R^{\{RECE\}}}_{f,genre}$	(31, 117)
$K_{plasmide}$	$ar{V}^{R^{\{plasmide\}}}_{f,genre}$	(273, 117)
K_{chr}	$ar{V}^{R^{\{chr\}}}_{f,genre}$	(546, 117)
$K_{monopartite}$	$ar{V}_{f,genre}^{\{monopartite\}}$	(530, 117)
$K_{multipartite}$	$ar{V}^{G\{monopartite\}}_{f,genre}$	(29, 117)
$K_{alphamono}$	$V_f^{G^{\{monopartite\}}_{alphaproteo}}$	(186,117)
$K_{alphamulti}$	$V_f^{G^{\{multipartite\}}_{alphaproteo}}$	(32,117)
$K_{betamono}$	$V_f^{G_{betaproteo}^{\{monopartite\}}}$	(93, 117)
$K_{betamulti}$	$V_{f}^{\{multipartite\}}$	(40, 117)

TABLE 6.5: Propriétés des classes utilisées pour l'analyse par régression logistique.

 a : Les notations utilisées sont celles introduites précédemment (§6.1 et §6.4.2), R représentant l'ensemble des réplicons.

 b : dimensions de la matrice associée aux données.

6.4.3 Résultats et discussion

Les régressions logistiques ont été réalisées pour des couples d'ensembles d'observations univariées et continues. Les OR sont tous calculés pour des observations ayant des attributs continus et indiquent l'augmentation de probabilité, selon le modèle estimé, qu'une observation appartienne à une classe donnée pour l'augmentation d'une unité de la valeur de l'attribut considéré. Par exemple, pour l'étude de la fonction ACLAME RepA $E\ B$ entre les classes $K_{betamono}$ et $K_{betamulti},$ un OR de 7,7 est estimé avec une P_{value} de 4, $1 * 10^{-9}$ (Table 6.6), ce qui signifie qu'un génome ayant 2 gènes codant pour des protéines annotées RepA, RepE, RepB dans ACLAME aura 7, 7 * 2 = 15, 4 fois plus de chances d'être un génome multipartite qu'un génome ne possédant aucun gène codant des protéines RepAEB. Pour l'étude entre K_{chr} et $K_{plasmide}$, un ensemble de réplicons normés par genre taxonomique où la moitié des réplicons possèdent des gènes codant des protéines annotées ParA, ParM dans ACLAME aura 0,5 * 0,4 = 0,2 fois plus de chance d'être un ensemble de chromosomes qu'un ensemble où aucun des réplicons ne possède de tels gènes. Ces modèles de régression reposent cependant sur l'hypothèse que les coefficients β_1 sont constants, ce qui peut s'avérer faux dans certains cas. Par exemple, la probabilité pour un réplicon d'acquérir un gène d'une source externe est vraisemblablement différente de celle qu'un gène existant se duplique.

TABLE 6.6: Résultats des régressions logistiques entre les différentes classes d'éléments génomiques définis (Table 6.5) pour chaque fonction des STIG.

 P_{value} : Signifiance des différents modèles (*i.e.*, probabilité que β_1 soit différent de 0). Modèle significatif (vert foncé) : $0 < P_{value} < 0.01$, peu significatifs (vert clair) : $0.01 < P_{value} < 0.05$ et non significatifs (non représentés) : $0.05 < P_{value}$. OR : odd-ratios des coefficients. Le code couleur indique l'importance de l'OR envers la première classe (tend vers l'orange) ou la seconde classe (tend vers le bleu).

K_{chr}/K_{plas}	smide	
Fonction	P_{value}	OR
kegg hupB	1.2e-53	97.6
kegg dnaG	2.1e-50	1861.5
kegg dnaA	2.5e-44 3.0e-44	2118.9
kegg dnaB	1.1e-43	1992.9
kegg ssb	1.7e-43 5.9e-41	298.3
kegg xerD	1.3e-38	26.6
kegg parA soj kegg ftsK spoIIIE	2.7e-38 2.8e-37	9.9
kegg E3.5.1.28B amiA	6.4e-36	46.4
amiB amiC kegg scpB	7.5e-32	102.5
kegg ftsZ	3.1e-31	2747.0
aclame Helicase kegg parC	1.6e-27 3.0e-27	71.1 4149.3
kegg cbpA	8.2e-27	2608.4
kegg par£ kegg fts£	7.3e-26 4.2e-24	2.3
kegg mreB	1.3e-21	1598.2
aclame DNAhelicase kegg dps	5.8e-21 9.1e-21	33.6 65.3
aclame AT-	2.2e-20	19.4
Pase.tyrK.exoP kegg iciA	7 1e-20	32
kegg lrp	1.6e-19	8.4
kegg minD kegg rob	3.1e-19 6.3e-19	42.8 5.3
kegg acrA	6.6e-19	2.8
kegg mrp kegg gidR remC	6.6e-17	2599.3
aclame RepA E B	1.7e-16	0.0
kegg ftsW spoVE	5.7e-16	4266.4
kegg diaA	0.9e-10 1.5e-15	81.9
kegg ftsQ	1.7e-15	2135.0
aclame PSK higBA kegg ihfB himD	3.3e-15 1.2e-14	3.4 68.4
kegg gidA mnmG	5.2e-13	1477.2
MTO1 kegg hfa	1.4e-12	121.7
kegg ihfA himA	1.7e-12	63.8
kegg rodA mrdB aclame ParB	2.8e-12 5.7e-12	1233.1 0.1
kegg dnaC	6.0e-12	2.6
kegg ftsX kegg ftsA	9.3e-12 9.5e-12	972.9
aclame PSK mazEF	1.2e-11	5.2
kegg scpA	1.4e-11	789.4
aclame XerTyrosine	2.9e-11 7.6e-11	2.0
aclame ParA.ParM	1.5e-10	0.4
vapBC.vag	1.2e-09	3.9
kegg fic	3.1e-09	10.3
kegg minC	3.8e-09 4.4e-09	52.3 172.3
kegg zapA	8.2e-09	602.8
kegg minE kegg ftsI	9.0e-09 9.8e-09	152.9 47.0
aclame RuvB	1.2e-08	433.0
kegg smc kegg mreD	1.6e-08 1.8e-08	3090.5
aclame PSK relBE	2.7e-08	3.5
aclame PSK parDE kegg sepF	5.5e-08	2.3
aclame FtsK.SpoIIIE	1.9e-07	6.0
aclame PSK phD.doc	3.2e-07	11.9
kegg hda	7.3e-07	149.1
kegg ftsB	1.1e-06	167.2
aclame serinerecombi-	1.5e-06 2.5e-06	1.4
nase kogg sul 4	3 20 00	17 5
kegg divIVA	4.1e-06	128.0
kegg hns	1.1e-05	2.8
kegg ftsL aclame PSK HicAB	1.2e-05 4.3e-05	91.5 25.2
kegg divIC divA	4.9e-05	90.5
kegg zipA kegg ftsN	7.9e-05 1.6e-04	66.0 53.0
aclame DNArepair	2.2e-04	34.0
kegg hupA aclame Tyrosinere-	2.7e-04 3.4e-04	15.1 3.3
cOrfA		
kegg seqA kegg mukB	1.6e-03 2.3e-03	25.9 27.4
kegg mukE	3.1e-03	21.0
kegg mukF	3.7e-03	19.6
aclame RepA	5.9e-03	0.7
kegg dnaB2 dnaB	6.7e-03	12.6
kegg ezrA	1.0e-02	13.7
aclame RepR S E	1.3e-02	0.0
aclame IrIA aclame primase LtrC	1.4e-02 3.1e-02	1.8
aclame Rop	3.2e-02	0.0
aciame RNApolyme- rase	3.2e-02	6.3
aclame PSK ccd	4.6e-02	3.9

K_{chr}/K_{RECE}						
Fonction	Pvalue	OR				
kegg dnaG	1.9e-21	205.3				
kegg dnaA kegg ffsZ	1.1e-19 1.2e-19	239.6				
kegg dnaB	5.1e-19	429.4				
kegg ssb	5.0e-18	160.6				
kegg parC	3.0e-16	134.0				
kegg ftsW spoVE kegg ftsI	4.4e-16 7.0e-16	87.7 3.9				
kegg gidB rsmG	2.2e-15	252.5				
kegg parE	5.7e-15	350.1				
kegg mrp aclame Helicase	1.3e-14 1.9e-13	35.0				
kegg cbpA	9.9e-13	22.8				
kegg mreB	3.9e-12	40.1				
kegg ftsQ	1.3e-11 2.0e-10	99.3				
amiB amiC	2.96-10	0.5				
kegg rodA mrdB	9.7e-10	33.0				
kegg ftsX	1.3e-08	13.8				
kegg mreC kegg ftsA	1.3e-08 2.2e-08	40.3 24.6				
kegg hupB	2.3e-08	6.7				
kegg ftsK spoIIIE	2.7e-08	15.8				
kegg gidA mnmG MTO1	2.9e-08	110.6				
kegg xerC	3.1e-08	8.8				
kegg xerD	4.1e-08	3.4				
aclame RuvB	5.7e-08	17.7				
kegg scpB kegg scpA	1.8e-07 5.7e-07	25.8				
kegg ftsE	1.3e-06	1.1				
aclame ParA.ParM	4.0e-06	0.3				
kegg zapA kegg parA soj	7.4e-06	17.3				
kegg smc	1.4e-05	131.9				
aclame ParB	1.4e-05	0.2				
kegg dps	3.5e-05	8.4				
kegg mreD kegg minD	0.8e-05 1.6e-04	2.3				
aclame RepA E B	1.9e-04	0.1				
aclame DNAhelicase	2.7e-04	4.1				
kegg htq kegg ihfB himD	3.0e-04 4.9e-04	6.9 8.4				
kegg diaA	1.2e-03	38.4				
kegg ihfA himA	1.4e-03	10.5				
aclame serinerecombi-	1.5e-03	2.9				
kegg parB spo0J	3.0e-03	2.1				
kegg fis	3.3e-03	7.9				
kegg trmFO gid kegg hda	4.4e-03 5.3e-03	8.3				
kegg ftsB	5.4e-03	16.1				
1 1. 1374	1.1 00	10.4				
kegg divIVA kegg slmA ttk	1.1e-02 1.2e-02	4.6				
kegg minC	1.2e-02	3.0				
kegg sepF	1.4e-02	12.3				
kegg acrA	1.7e-02 2.4e-02	1.1				
kegg ftsL	2.4e-02 2.7e-02	9.8				
aclame CopG	2.7e-02	0.2				
aclame PSK mazEF	2.9e-02	2.6				
kegg dam	3.6e-02	2.0				
aclame cdc6archee	4.4e-02	0.1				
aclame DNAbinding	4.4e-02	0.1				
aclame RepH aclame PSK parC	4.4e-02 4.4e-02	0.1				
aclame RepA BCopB	4.4e-02	0.1				
aclame cdsD	4.4e-02	0.1				
aclame plasmidmainte-	4.4e-02	0.1				
aclame Rop	4.4e-02	0.1				
kegg divIC divA	4.7e-02	8.1				
aclame RepR S E	4.9e-02	0.1				

$K_{RECE}/K_{plasmide}$						
Fonction	P_{value}	OR				
kegg ftsE	4.0e-11	1.9				
kegg lrp	5.4e-11 5.4e-11	8.1				
kegg acrA	5.3e-10	2.7				
kegg rob	3.4e-08	4.2				
kegg ftsI	2.2e-07	76.7				
kegg nupB	2.4e-07 4.5e=07	11.0				
kegg cbpA	5.6e-07	36.1				
aclame AT- Pase tyrK exoP	8.5e-07	9.3				
kegg parB spo0J	2.3e-06	4.1				
kegg xerD	2.5e-06	6.2				
kegg parA soj	8.4e-06	3.8				
aclame KuvB aclame PSK	1.4e-05	37.8 5.8				
vapBC.vag						
kegg mreB	1.4e-05	24.1				
kegg mrp	2.5e-05	86.2				
kegg minC kegg minE	5.8e-05	76.8				
aclame PSK parDE	7.8e-05	3.4				
aclame FtsK.SpoIIIE	9.9e-05	9.8				
aclame Helicase	1.1e-04	4.6				
aclame DNAhelicase	1.3e-04	9.8				
kegg xerC kegg parE	1.8e-04 2.4e=04	15.8				
kegg hns	3.8e-04	2.8				
kegg parC	4.6e-04	12.3				
kegg ftsX	4.8e-04	146.2				
aclame DNArepair	5.7e-04 6.1e-04	43.6				
aclame Tyrosinere-	7.4e-04	8.7				
cOrfA						
kegg hfq	8.1e-04	19.3				
kegg ftsW spoVE kegg fts7	8.2e-04	55.0 16 5				
aclame PSK higBA	9.7e-04 1.2e-03	2.5				
kegg ftsA	2.5e-03	41.7				
kegg dnaG	2.5e-03	4.5				
kegg rodA mrdB	2.6e-03	55.3				
kegg dnaA	2.9e-03 3.5e-03	8.3				
aclame CopG	4.6e-03	23.1				
kegg E3.5.1.28B amiA	4.6e-03	3.0				
aclame PSK HicAB	4.8e-03	15.1				
aclame XerTyrosine	6.3e-03	1.6				
kegg zapA	7.3e-03	56.1				
kegg dnaB kogg ihfR himD	8.2e-03	3.7				
kegg fic	8.6e-03	7.2				
kegg dps	8.7e-03	6.7				
kegg mreC	8.9e-03	32.8				
kegg gidB rsmG	9.0e-03	32.2				
aclame RepC	9.6e-03	20.0				
kegg fis	1.4e-02	25.1				
kegg ftsK spolllE	1.4e-02	4.2				
kegg slmA ttk	1.5e-02	7.5				
aclame serinerecombi-	1.8e-02	0.4				
nase kegg mukF	1.0e.09	18.2				
kegg mukB	1.9e-02	18.2				
kegg mukE	1.9e-02	18.2				
kegg mreD	1.9e-02	18.2				
kegg trmFO gid	1.9e-02	18.0				
kegg scpA	2.1e-02	16.6				
kegg dam	2.4e-02	4.3				
kegg gidA mnmG	4.3e-02	18.2				
MTO1 kegg dnaC	4.6e-02	15				
kegg ihfA himA	4.9e-02	6.9				
_0						

Légende

$e^4 < OR$
$e^3 < OR < e^4$
$e^2 < OR < e^3$
$e^1 < OR < e^2$
$e^0 < OR < e^1$
$e^{-1} < OR < e^0$
$e^{-2} < OR < e^{-1}$
$e^{-3} < OR < e^{-2}$
$e^{-4} < OR < e^{-3}$
$OR < e^{-4}$

Les résultats des différentes études par régression logistique de la pertinence des fonctions des STIG pour les différentes classes de réplicons sont présentées Tables 6.6 et 6.7.

- Les chromosomes se distinguent très nettement des plasmides sur la majorité des fonctions STIG (Table 6.6). Seules les annotations ACLAME : Rep, ParA et ParB, Rop et TrfA sont caractéristiques du type *plasmide*. Certaines annotations comme DnaA, DnaB ou FtsZ semblent être beaucoup plus spécifiques des chromosomes que d'autres, comme par exemple, FtsE, DnaC, H-NS... Ces résultats décrivent la séparation nette entre chromosomes et plasmides du point de vue des STIG observés dans les précédentes analyses.
- Le biais de distribution des gènes des STIG entre chromosomes et RECE différe de celui existant entre chromosomes et plasmides. De nombreuses fonctions montrant un biais de présence entre chromosomes et plasmides ne sont pas ou peu retrouvées entre chromosomes et RECE. Certaines annotations de type "chromosomique" ne sont pas exclues des RECE : DnaG, DnaA, FtsZ... Par contre, les RECE se distinguent des chromosomes par des annotations ACLAME propres, témoignant d'un lien fort d'au moins une partie des RECE avec les plasmides.
- Les biais observés entre RECE et plasmides indiquent des spécificités fortes des STIG des RECE par rapport aux plasmides classiques. Une des difficultés est d'identifier des différences RECE/plasmides observées pour un groupe taxonomique donné, témoignant d'un événement génomique ponctuel (THG ou duplication) et de les différencier de biais globaux, observés pour des RECE issus de groupes taxonomiques éloignés. Les protéines annotées "Hfq" des Bêtaprotéobactéries où un gène hfq supplémentaire est présent chez les Burkholdériales à génome multipartite (à l'exception de Variovorax) illustre un cas de duplication génique probable, alors que le transfert de l'opéron minCDE du chromosome au RECE et à certains mégaplasmides des Rhizobiales représente un cas de THG [Slater et al., 2009]. Ces biais ponctuels, spécifiques à certains groupes taxonomiques de génomes, sont caractérisés par des OR extrêmes (fort ou faible), qui reflètent la non-conformité des réplicons. Parmi les biais trouvés pour les différents groupes taxonomiques, aucun ne semble universel à l'ensemble des RECE. Certaines tendances sont cependant très marquées, comme les présences de gènes codant pour des régulateurs de type AcrA, Lrp, IciA, Rob. Les gènes codant les constituants centraux du système de partition : ParA/ParB, MinD, se retrouvent aussi de façon significative sur les RECE. Ces traits sont confirmés dans la comparaison entre génomes monopartites et multipartites.
- Aucun biais global n'est détectable entre génomes mono- et multipartites. Contrairement à la comparaison chromosome/plasmide (présence de gènes *dnaA*, *ftsZ*, *parC*...), aucune caractéristique forte ne permet d'identifier clairement un génome multipartite. Cependant, des tendances caractéristiques des génomes multipartites apparaissent (Table 6.7).

 $COR < \epsilon$

OR <

TABLE 6.7: Résultats des régressions logistiques entre les différentes classes de génomes définis (Table 6.5) pour chaque fonction des STIG.

 P_{value} : Signifiance des différents modèles (*i.e.*, probabilité que β_1 soit différent de 0). Modèle significatif $(\text{vert foncé}) : 0 < P_{value} \neq 0.01, \text{ peu significatifs (vert clair)} : 0.01 < P_{value} < 0.05 \text{ et non significatifs (non représentés)} : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coefficients. Le code couleur indique l'importance (non représentés) : 0.05 < P_{value}. OR : odd-ratios des coeffi$ de l'OR envers la première classe (tend vers l'orange) ou la seconde classe (tend vers le bleu).

$K_{multipartite}/K_{monopartite}$			$K_{betamulti}/K_b$	$K_{alphamulti}/K_{alphamond}$			
Fonction	P_{value}	OR	Fonction	P_{value}	OR	Fonction	Pvalue
aclame ParB	6.8e-07	2.0	kegg hfq	2.5e-12	171.3	kegg xerC	3.2e-07
kegg iciA	2.2e-06	1.0	kegg hns	1.5e-09	1.9	kegg minE	6.6e-06
kegg lrp	4.1e-06	1.2	kegg xerC	2.4e-09	20.3	aclame FtsK.SpoIIIE	3.8e-05
aclame ParA.ParM	1.4e-05	1.7	kegg minD	4.5e-09	4.8	kegg lrp	7.4e-05
aclame RepC	7.1e-05	1.6	kegg iciA	6.5e-09	1.1	kegg smc	1.6e-04
kegg parB spo0J	1.2e-04	1.4	kegg ftsI	8.7e-09	3.2	kegg mrp	1.7e-04
aclame PSK parDE	1.5e-04	1.3	aclame RepA E B	4.8e-08	17.9	kegg minC	2.7e-04
aclame PLdimerreso-	2.8e-04	2.1	kegg parA soj	5.0e-08	1.9	kegg rodA mrdB	4.4e-04
lution			kegg lrp	5.8e-08	1.4	kegg mreB	5.4e-04
aclame FtsK.SpoIIIE	3.3e-04	2.5	kegg rob	7.9e-08	1.5	aclame ParA.ParM	6.2e-04
kegg hns	3.7e-04	1.4	aclame AT-	3.3e-07	1.9	kegg scpA	7.3e-04
aclame PSK	4.3e-04	1.3	Pase.tyrK.exoP			kegg mreC	1.2e-03
vapBC.vag			kegg ftsK spoIIIE	4.9e-07	14.7	aclame ParB	1.2e-03
aclame DnaB	4.5e-04	7.6	kegg acrA	7.0e-07	1.2	kegg fis	2.1e-03
kegg parA soj	5.3e-04	1.3	kegg sulA	1.3e-06	5.6	kegg hfq	3.8e-03
kegg minE	1.1e-03	3.8	kegg ftsX	1.5e-06	0.0	kegg ftsE	3.8e-03
aclame AT-	2.0e-03	1.3	kegg parB spo0J	3.4e-06	1.6	kegg ftsX	5.1e-03
Pase.tyrK.exoP			kegg ftsN	5.1e-06	7.3	aclame RepA E B	5.2e-03
kegg ftsE	2.4e-03	1.0	kegg mreB	8.4e-06	2.8	aclame AT-	6.0e-03
aclame RepA E B	2.8e-03	1.7	kegg ftsE	8.9e-06	1.1	Pase.tyrK.exoP	
aclame CopG	4.8e-03	4.6	kegg E3.5.1.28B amiA	5.4e-04	0.4	kegg scpB	6.9e-03
kegg minC	8.1e-03	2.5	amiB amiC				
kegg hfq	9.1e-03	2.2	aclame ParA.ParM	8.0e-04	1.8	kegg cbpA	1.2e-02
	4.0.00		kegg dps	1.4e-03	3.8	aclame RepC	1.3e-02
kegg hupB	1.2e-02	1.4	kegg dnaG	2.0e-03	4.7	kegg hns	1.4e-02
kegg ftsN	1.2e-02	2.4	kegg fic	2.8e-03	0.2	kegg parA soj	1.6e-02
aclame DNArepair	1.3e-02	1.9	kegg hupB	3.2e-03	1.7	kegg iciA	1.9e-02
kegg sulA	1.3e-02	2.2	aclame XerTyrosine	4.0e-03	1.1	aclame PSK HicAB	3.1e-02
kegg diviva	1.4e-02	0.2	aclame PSK higBA	5.4e-03	1.2	aclame helicase	3.5e-02
aclame PSK epsi-	1.6e-02	3.1	aclame PSK	6.2e-03	1.2	kegg diaA	4.1e-02
lon.zeta	1.0.00	1.4	vapBC.vag			aclame RuvB	4.5e-02
acianie KepA	1.66-02	1.4	aclame Helicase	9.3e-03	1.4		
aciame nencase	2.16-02	3.9	1 1 4	1 1 00	0.0		
Regg seqA	2.16-02	3.1	kegg diaA	1.1e-02	2.8	т /	
acianie Hencase	2.86-02	1.0	aclame AerD	1.3e-02	3.4	Legen	de
kegg itsi	3.16-02	1.2	kegg rodA mrdB	1.5e-02	12.4		
kegg Itsz	3.20-02	1.9	aciame Fis	1.7e-02	2.1		
kegg acrA	3.36-02	1.1	kegg mreD	1.9e-02	11.0	$e^4 < C$	R
kegg mukr	3.4e-02 4.9= 09	3.0	kegg simA ttk	2.0e-02	1.8	3.00	. 4
aciame cucoarcnee	4.26-02	18.3	aciame DNArepair	2.0e-02	2.5	$e^{s} < OR$	$< e^{-}$
acianie nepH	4.20-02	18.3	kegg cDpA	2.4e-02	0.5	$c^2 < OB$	< 3
adame DNAbind	4.20-02	18.3	kegg part	3.4e-02	9.1	e < Oh	< e
adame plasmidmeint	4.30-02	18.0	kegg park	3.4e-02	9.1	$e^1 < OR$	$< e^2$
nance PSK	4.56-02	18.0	kegg zipA	3.8e-02	0.1		, 1
kegg mukE	4.3e-02	2.9				$e^{\circ} < OR$	$< e^{\perp}$
kegg cbpA	4.7e-02	1.2				$e^{-1} < OF$	$2 \leq e^0$
00 ·····							$\sim C$

La présence accrue de certains gènes apparaît corrélée avec l'architecture "multipartite" d'un génome :

- des gènes régulateurs, annotés *iciA*, *lrp*,
- des gènes impliqués dans la partition (*parA/parB,minC*),
- des gènes impliqués dans l'initiation de la réplication (*rep*), codant des initiateurs plasmidiques de type DnaB.
- des gènes impliqués dans la structure du génome (hns).

Certains gènes tels que hfq et xerC sont en plus grand nombre dans les génomes multipartites vs. monopartites chez les Alpha- et Bêtaprotéobactéries, confirmant l'existance de spécificité taxonomique.

kegg cbpA aclame XerD

5.0e-02 2.1

6.4.4 Conclusion

Dans l'analyse des biais de présence des gènes des STIG, il est difficile de distinguer les tendances témoignant d'une adaptation locale dans une lignée bactérienne donnée (par exemple les gènes repC présents chez les RECE des Rhizobiales et témoins de la présence de l'opéron repABC, caractéristique des mégaplasmides de ce groupe [Slater et al., 2009]), de caractéristiques globales partagées par l'ensemble des génomes multipartites et diagnostiques de cet état multipartite. De nombreux bruits peuvent perturber l'interprétation des résultats.

- Le problème majeur est le nombre réduit de données actuellement disponibles et la distribution spécifique des génomes multipartites dans certaines lignées (seulement 29 genres taxonomiques représentés pour $K_{multipartite}$), les classes K_{RECE} et $K_{multipartite}$ ne représentant que 5 à 10% des classes avec lesquelles elles sont comparées. Aucune stratégie particulière n'a été suivie afin de corriger les biais potentiels dus à cette disparité (l'état multipartite peut être qualifié d'"événement rare") [King and Zeng, 2001].
- Un autre problème vient de l'absence potentielle de variables explicatives. Toutes les fonctions des STIG peuvent ne pas être représentées (probablement) ou l'être incorrectement pour tout ou une partie des génomes, ce qui peut conduire à un manque d'information pour des génomes peu représentés et/ou étudiés. Les cyanobactéries en sont un exemple.
- Il est également possible que les attributs significativement biaisés soient corrélés à des facteurs externes autres que le type de réplicon ou l'état mono/multipartite des génomes (par exemple, l'écologie des espèces). À la question de savoir si le nombre de données est suffisant pour construire un modèle de régression robuste, certains auteurs [Hosmer Jr et al., 2013] recommandent que le nombre de paramètres p à inclure dans le modèle ne soit pas supérieur à :

$$p+1 \le \frac{\min\{|E_{True}|, |E_{False}|\}}{10} \tag{6.12}$$

Les différentes régressions conduites dans cette étude étant toujours monovariées (p = 1), ce critère est toujours respecté. Cependant compte tenu des nombreuses sources potentielles de bruits évoqués, les modèles estimés sont probablement, pour la plupart, des approximations de la réalité. Les OR doivent alors être interprétés avec précaution (surtout pour des P_{value} supérieures à 10^{-3}).

- Enfin, des corrélations entre variables ne sont pas prises en compte dans l'analyse : la présence d'un gène parA est très souvent couplée à la présence d'un gène parB, de même pour dnaA/dnaB, ftsK/ftsZ, ftsZ/ftsX... (existence d'opérons notamment), l'objectif principal de cette étude étant de donner une première estimation des spécificités fonctionnelles des différentes classes d'éléments génomiques.

Chapitre 7

Classification supervisée

Les résultats précédents de cette thèse ont montré l'existence de biais de distribution des gènes liés aux STIG selon le type de réplicons : chromosome, plasmide ou RECE. Cependant, pour certains réplicons extrachromosomiques, le statut d'essentialité pour l'hôte n'est pas clair et certains réplicons annotés "plasmide" ont eu leur statut récemment révisé et sont désormais considérés comme essentiels pour leur hôte [Landeta et al., 2011]. En se servant des annotations fonctionnelles comme attributs des réplicons, des analyses par classification supervisée sont donc conduites afin d'identifier des RECE potentiels parmi les plasmides. Ces analyses permettent, de plus, d'identifier des RECE se classant parmi les chromosomes, et inversement.

7.1 Jeux de données

Plusieurs jeux d'apprentissage (*training sets*) sont formés selon les annotations de RefSeq des réplicons : "chromosome", "plasmide" et "RECE", et des génomes : mono- ou multipartites. Les données sont de plus normées par genre pour éviter une sur-représentation de certains groupes taxonomiques.

Soit E_{chr} , $E_{plasmide}$, E_{RECE} , $E_{monopartite}$ et $E_{multipartite}$ les différents training sets, définis par :

- $\mathbf{E_{chr}}$ L'ensemble des chromosomes normé par le genre de l'hôte. Ainsi $E_{chr} = K_{chr}$, avec K_{chr} défini Table 6.5.
- $\mathbf{E}_{\mathbf{RECE}}$ L'ensemble des RECE normé par le genre de l'hôte. Ainsi $E_{RECE} = K_{RECE}$, avec K_{RECE} défini Table 6.5.
- $\mathbf{E}_{\mathbf{plasmide}}$ Pour ne pas inclure des plasmides pouvant être de potentiels RECE, les résultats du clustering par INFOMAP (§5.4.2) ont été utilisés. Les plasmides au sein des clusters de la procédure $f_{INFOMAP}(V^R)$ où sont également présents des réplicons de type "RECE" ou "chromosome" ne sont pas pris en compte.

Soit $Cl_{INFOMAP} = f_{INFOMAP}(V^R)$, le résultat du clustering par INFOMAP. $E_{plasmide}$ est défini par :

$$E_{plasmide} = \bar{V}_{f,genre}^{R_{classif}^{\{plasmide\}}}$$
(7.1)

où $R^{\{plasmide\}}_{classif}$ est défini par :

$$R_{classif}^{\{plasmide\}} = \bigcup_{\substack{C \in Cl_{INFOMAP} \\ C \cap R^{\{chr, RECE\}} = \emptyset}} C$$
(7.2)

- $\mathbf{E}_{\text{monopartite}}$ L'ensemble des génomes monopartites ayant un seul réplicon essentiel (chromosome) normé par le genre de l'hôte. Ainsi $E_{monopartite} = K_{monopartite}$, avec $K_{monopartite}$ défini Table 6.5.
- $\mathbf{E}_{\text{multipartite}}$ L'ensemble des génomes multipartites normé par le genre de l'hôte. Ainsi $E_{multipartite} = K_{multipartite}$, avec $K_{multipartite}$ défini Table 6.5.

 TABLE 7.1: Taille des ensembles formant les training sets utilisés pour les classifications supervisées.

Ensemble	E_{chr}	E_{RECE}	$R^{\{plasmide\}}_{classif}$	$E_{plasmide}$	$E_{monopartite}$	$E_{multipartite}$
Taille	548	31	2744	262	530	29

7.2 Algorithmes de classification

Plusieurs approches méthodologiques de classification supervisée sont possibles. Les différents algorithmes utilisés dans notre étude sont ici brièvement présentés :

TABLE 7.2: Principaux algorithmes de classification supervisée utilisés.

Logistic regression	La fonction $f_{reg}^{E_{training}}$ définie dans l'éq. 6.9 peut être utilisée pour classer des observations $o \notin E_{training}$: si $f_{reg}^{E_{training}}(o) > 0.5$ alors $o \in E_{True}$ [Hosmer Jr et al., 2013].
Arbres de	[Breiman et al., 1984] Le principe de ces algorithmes est de bâtir, en fonc-
décision	tion d'un jeu de données, un arbre de décision effectuant une série de tests
(AD)	séquentiels pour une observation et aboutissant à l'attribution d'une classe pour cette observation. Succinctement, le principe est, pour la création de chaque nœud (ou embranchement), de choisir le couple Attribut/Valeur-seuil permettant de séparer au mieux le <i>training set</i> . Les différentes stratégies per- mettant la conception d'arbres de décision sont détaillées dans [Izenman, 2008, p.281]. Les paramètres de ces algorithmes sont, entre autres, le nombre maxi- mal d'embranchements successifs de l'arbre, les modalités de création des nœuds de l'arbre (nombre minimum d'observations par embranchement, par coupure) et les processus de raffinage (<i>Tree pruning</i>).

Random
forest (RF)[Breiman, 2001] Cette procédure fait partie des méta-classifieurs introduits
 $\S3.4.7.1$. Le principe est de créer une forêt d'arbres dont chacun est réalisé
par un échantillon de bootstrap effectué sur les observations de $E_{training}$, si-
milairement au principe du bagging. Les arbres construits possèdent une part
supplémentaire de hasard en choisissant aléatoirement, pour la construction de
chaque noeud, seulement un sous-ensemble d'attributs [Izenman, 2008, p.537].
Cette procédure a pour effet de réduire la variance ainsi que le biais des
résultats obtenus par un arbre de décision seul. Les paramètres sont, en plus
de ceux propres aux arbres de décision, le nombre d'arbres de la forêt et la
taille des sous-ensembles d'attributs.

Extremely randomized trees (ERT) [Geurts et al., 2006] Cet algorithme est similaire à celui des *Random Forest* mais utilise une part d'aléatoire encore plus importante. Pour chaque attribut testé lors de la création d'un nœud, la valeur seuil est tirée de façon aléatoire. Cela a pour effet de réduire davantage la variance mais augmente légèrement le biais de classification (ou taux d'erreur, *cf.* notion de décomposition *biaisvariance* [Witten et al., 2011, p. 354]).

Naive Bayes classification

[Larose, 2006, p.219] Soit une observation o tirée d'un jeu de données $E = E_1 \cup ... \cup E_k$ organisé en k classes. En supposant que les attributs de o sont indépendants conditionnellement, on peut estimer, à partir du théorème de Bayes, que la probabilité que o appartienne à E_i est proportionnelle à :

$$P(o \in E_j) \propto \prod_{i=1}^{|v_o|} P(v_o[i]|o \in E_j).P(E_j)$$

$$(7.3)$$

les probabilités $P(E_j)$ (estimée par le ratio $\frac{|E_j|}{|E|}$) et $P(v_o[i]|o \in E_j)$ (estimée, par exemple, par la proportion du nombre d'observations o' dans E_j telles que $v_{o'}[i] = v_o[i]$) pouvant être facilement estimées à partir des données. o appartient alors à la classe E_j pour laquelle $P(o \in E_j)$ est maximale. Les paramètres à considérer sont la spécification de potentielles relations de dépendance entre les attributs.

Support Vector Machine (SVM) [Cortes and Vapnik, 1995] Ces ensembles d'algorithmes d'apprentissage supervisé ont pour objectif de définir un ou plusieurs hyperplans au sein d'un espace où sont représentées les observations, de façon à séparer au mieux les différentes classes d'observations. Ces hyperplans sont définis par rapport aux vecteurs supports, représentant les observations les plus proches des limites théoriques entre les classes. Un deuxième aspect important des SVM est que, dans le cas où les observations ne sont pas séparables linéairement dans l'espace de représentation, une fonction "noyau" f_k peut être utilisée afin de projeter les observations dans un espace de plus grande dimension où elles seront potentiellement séparables. Un exemple simple de f_k est la fonction polynomiale : $f_k(x, y) = (\langle x, y \rangle + c)^d$, x et y étant deux vecteurs d'observations. Le principe général et la méthodologie des SVM sont très bien détaillés dans [Hamel, 2011]. Les paramètres à définir sont, entre autres, f_k , les différents paramètres propres à f_k (c et d...) et le coût Co de la pénalisation des observations incorrectement séparées par l'hyperplan. Lors de l'utilisation des algorithmes de type Ensemble (Random Forest et Extremely randomized trees), un estimateur de performance, l'out-of-bag estimate (ou OOB_{score}), peut être calculé à partir des différents tirages de bootstrap de chaque classifieur des procédures : pour chaque tirage de bootstrap, une partie des observations (ou out-of-bag (OOB)) ne sont pas choisies pour la construction de $E_{training}$. Ces observations peuvent ainsi servir de jeu de données test pour mesurer l'efficacité des classifieurs [Izenman, 2008, p.507]. À partir des observations des OOB, les importances des variables dans la classification des observations peuvent être calculées, en permutant une à une la valeur des attributs et en comparant le nouveau OOB_{score} à celui obtenu sans permutation [Breiman, 2001][Izenman, 2008, p.543]. Enfin, différentes méthodes peuvent servir à calculer la probabilité qu'une observation appartienne à une certaine classe selon, par exemple, la distance à laquelle celle-ci se trouve de l'hyperplan (SVM) [Rüping, 2004]. Pour les méthodes de type Ensemble, une façon simple d'estimer ces probabilités est de considérer, pour une observation donnée, le rapport : nombre de classements dans la classifieurs la procédure.

7.3 Procédures

7.3.1 Classification des plasmides

Afin de détecter des réplicons annotés "plasmide" mais susceptibles d'être des RECE, l'ensemble $V_f^{R\{plasmide, RECE\}}$ des réplicons extra-chromosomiques est classé en utilisant $E_{learning} = \{E_{RECE}, E_{plasmide}\}$ comme *learning-set*. La différence de taille entre les deux sous-ensembles (Table 7.1) peut entraîner un déséquilibre pour les résultats apportés par les classifieurs [Han et al., 2012, p.385] et grandement favoriser la classe "plasmide" lors de la classification d'un réplicon.

Similairement à la démarche proposée par [Larose, 2006, p.213], une seconde procédure de classification incluant une procédure d'échantillonnage aléatoire sans remise sur $R^{\{plasmide\}}_{classif}$ est alors conduite. Les étapes consistent à :

- 1 Tirer un échantillon R_{ech} de même taille que $R^{\{RECE\}}$ avec remise de $R^{\{plasmide\}}$ et le normer par genre : $E_{ech} = \bar{V}_{f,genre}^{R_{ech}}$.
- 2 Effectuer la procédure de classification : $f_{classif}^{E_{training}}(V_{f}^{R^{\{plasmide,RECE\}}})$, sur l'ensemble des réplicons plasmidiques et RECE décrits fonctionnellement par : $V_{f}^{R^{\{plasmide,RECE\}}}$ avec $E_{training} = \{E_{chr}, E_{ech}\}$.
- **3** Effectuer les étapes 1 et 2 n fois.
- 4 Établir la moyenne des différents indices calculés pour les n classifieurs : probabilités d'appartenir à une classe pour une observation donnée, scores obtenus pour les procédures de cross-validation, scores $OOB_{estimate}$ obtenus pour les procédures de Random Forest et Extremely randomized trees, et scores d'importance pour les attributs. On peut alors estimer qu'une observation appartient à une classe donnée ("plasmide" ou "RECE") si la majorité des classifieurs l'ont attribuée à cette classe.

Le $E_{learning}$ de cette procédure d'échantillonnage aléatoire sans remise est alors désigné par $\{E_{RECE}, E_{plasmide}\}^{it}$. En omettant une partie des observations de $E_{plasmide}$ à chaque itération, cette procédure fait baisser la *précision* du classifieur en augmentant le taux, FP, de faux-positifs détectés ("vrai" plasmides annotés "RECE"). Elle augmente aussi vraisemblablement la variance des résultats du classifieur en rajoutant une part d'aléatoire : le tirage dans $E_{plasmide}$. Cependant, en incluant un plus grand nombre de probables vrais RECE classés comme TP, cette procédure réduit le taux, FN, de faux-négatifs de l'analyse et accroît alors la sensibilité du classifieur (cf. éq. 3.9).

7.3.2 Classification des chromosomes

La classification des chromosomes a pour objectif d'étudier de potentiels chromosomes classés parmi les plasmides, ou plus généralement les réplicons extra-chromosomiques, et inversement s'il existe des réplicons extra-chromosomiques se classant parmi les chromosomes. Pour cela, deux procédures de classification $f_{classif}^{E_{training}}(V_f^{R^{\{plasmide,RECE,chr\}}})$ ont été réalisées en utilisant d'une part $E_{training} = \{E_{chr}, E_{plasmide}\}$, et d'autre part $E_{training} = \{E_{chr}, E_{RECE}\}^{it}$ qui, similairement à la classification des plasmides, désigne une procédure d'échantillonnage aléatoire sans remise, effectuée sur $R^{\{chr\}}$.

7.3.3 Classification des génomes

Malgré les faibles biais détectés entre génomes multi- et monopartites dans l'analyse par régression logistique (Table 6.6), les génomes de $\bar{V}_{f,genre}^G$ sont classés en utilisant $E_{training} = \{E_{multipartite}, E_{monopartite}\}^{it}$ qui, similairement à la classification des plasmides désigne une procédure d'échantillonnage aléatoire sans remise effectuée sur $G^{\{monopartite\}}$. L'objectif est alors de détecter de potentiels génomes multipartites considérés comme monopartites jusqu'à présent.

7.4 Sélection des algorithmes et des paramètres

Afin d'identifier l'approche méthodologique optimale, les différents algorithmes de classification supervisée ont été appliqués à ces jeux de données (§7.2). Pour chaque classification, une procédure de Cross-Validation (CV) est réalisée par Stratified 10-folds, suivant les recommandations de [Han et al., 2012, p.370]. $E_{training}$ est séparé en 10 partitions équilibrées par rapport aux proportions relatives des classes de $E_{training}$. Chaque partition sert ensuite de test set par rapport aux neuf autres. Pour une partition $K = \{K_1, ..., K_k\}$ par K-Fold d'un training set, le CV_{score} de la procédure désigne le pourcentage moyen de classification correcte, donnée par la procédure de CV pour chaque partition K_i [Hamel, 2011, chap.9] :

$$CV_{score} = 1 - CV_{erreur} \tag{7.4}$$

avec :

$$CV_{erreur} = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{|K_i|} Err(f_{classif}^{K \setminus K_i}(K_i))$$
(7.5)

où $Err(f_{classif}^{K \setminus K_i}(K_i))$ désigne le nombre d'observations de K_i classées incorrectement par $f_{classif}^{K \setminus K_i}$ et $K \setminus K_i = \{k | k \in K \land k \notin K_i\}.$

Pour les méthodes de type Ensemble, l' OOB_{score} est aussi reporté. Pour les procédures de classification incluant des échantillonnages aléatoires, CV_{score} et OOB_{score} désignent la moyenne des CV_{score} et des OOB_{score} obtenus pour chaque itération. Pour ces procédures, les écart-types $\sigma_{CV_{score}}$ et $\sigma_{OOB_{score}}$ sont aussi calculés.

Les paramètres de SVM (kernel polynomial d'ordre 2 et C = 10) sont choisis selon les CV_{score} des classifications plasmide/RECE. De même, les paramètres par défaut des arbres de décision ont été utilisés car ils donnent de meilleurs CV_{score} . Le nombre d'arbres dans les procédures RF et ERT est fixé à 1000. Le nombre d'itérations n des procédures d'échantillonnage est fixé à 100.

Les résultats obtenus sont très similaires sur les training set $\{E_{RECE}, E_{plasmide}\}$, à l'exception du classifieur Naive (Table 7.3). Ces bonnes performances peuvent s'expliquer par la relative grande taille de $E_{plasmide}$ par rapport à E_{RECE} et la facilité à classer les réplicons de type "plasmide" parmi les plasmides. Ainsi, les faibles valeurs de CV_{score} sont dues aux faibles valeurs des $Err(f_{classif}^{K\setminus K_i}(K_i))$ grâce à la large proportion d'observations "plasmide" présentes dans les K_i . Les résultats obtenus sur $\{E_{RECE}, E_{plasmide}\}^{it}$, bien qu'ayant des proportions comparables d'observations "plasmide" et "RECE", témoignent de la plus grande difficulté à classer les réplicons annotés "RECE" dans la classe des RECE. Les scores obtenus pour les training sets $\{E_{RECE}, E_{chr}\}^{it}$ et $\{E_{RECE}, E_{plasmide}\}^{it}$ sont ainsi des meilleurs indicateurs de performance pour la classification des RECE.

Sur ces training sets, les classifieurs RF et ERT donnent les résultats les plus performants par comparaison aux résultats de SVM, Naive et AD. Cette constatation est confirmée par le résultat d'un test de Kolmogorov-Smirnov entre les distributions de CV_{score} de SVM et de RF et ERT, respectivement ($p_{value} << 0.05$). Les résultats de RF et ERT ne sont cependant pas significativement différents ($p_{value} > 0.05$). Enfin, les résultats de la classification de { $E_{monopartite}, E_{multipartite}$ }^{it} mettent en évidence une faible efficacité des classifieurs dans la séparation génome multi/monopartite.

${f E_{training}}$	Classifieur	CV_{score}	$\sigma_{CV_{score}}$	OOB_{score}	$\sigma_{OOB_{score}}$
$\{E_{RECE}, E_{plasmide}\}$	\mathbf{SVM}	0.96	-	-	-
	Naive	0.90	-	-	-
	\mathbf{AD}	0.96	-	-	-
	\mathbf{RF}	0.96	-	0.96	-
	\mathbf{ERT}	0.96	-	0.96	-
$\{E_{RECE}, E_{plasmide}\}^{it}$	\mathbf{SVM}	0.90	0.03	-	-
	Naive	0.83	0.03	-	-
	\mathbf{AD}	0.85	0.04	-	-
	\mathbf{RF}	0.93	0.02	0.93	0.02
	\mathbf{ERT}	0.92	0.02	0.93	0.02
$\overline{\{E_{chr}, E_{plasmide}\}}$	\mathbf{SVM}	1.0	-	-	-
-	Naive	0.98	-	-	-
	\mathbf{AD}	1.0	-	-	-
	\mathbf{RF}	1.0	-	1.0	-
	\mathbf{ERT}	1.0	-	1.0	-
$\overline{\{E_{RECE}, E_{chr}\}^{it}}$	\mathbf{SVM}	0.93	0.02	-	-
	Naive	0.87	0.01	-	-
	\mathbf{AD}	0.95	0.02	-	-
	\mathbf{RF}	0.98	0.01	0.98	0.01
	\mathbf{ERT}	0.98	0	0.98	0.01
$\{E_{monopartite}, E_{multipartite}\}^{it}$	\mathbf{SVM}	0.72	0.04	-	-
	Naive	0.59	0.04	-	-
	\mathbf{AD}	0.68	0.04	-	-
	\mathbf{RF}	0.77	0.03	0.77	0.03
	\mathbf{ERT}	0.78	0.03	0.78	0.03

 TABLE 7.3: Comparaison des classifieurs pour la classification supervisée des réplicons.

SVM, Naive Bayes classification (**Naive**), Arbre de décision (**AD**) Random Forest (**RF**), Extra Randomised Trees (**ERT**). Training sets : $\{E_{RECE}, E_{chr}\}^{it}$ et $\{E_{RECE}, E_{plasmide}\}^{it}$ désignent les procédures itératives décrites (cf. §7.3).

7.5 Logiciels utilisés

- Scikit-learn Les classifieurs SVC, GaussianNB, DecisionTreeClassifier, RandomForest-Classifier, ExtraTreesClassifier de la librairie Python, Scikit-learn ont été utilisés. Les probabilités des classes pour les observations, les OOB_{score} , et les importances des attributs ont été calculées avec les fonctions $predict_proba$, $oob_score_$, feature_importances_, pour chaque procédure RF et ERT.
- **Python** a été utilisé pour la réalisation des *pipelines* analytiques, du traitement des données et des procédures CV.
- ${\bf R}\,$ a été utilisé pour faire les tests d'hypothèses (Kolmogorov-Smirnov).

	(Λ) [EDROP E ,]	
Actinobactéries	(A) $\{ LRECE, Lplasmide \}$	
Actinomycétales		
NC_016113	S. cattleya NRRL 8057	0.727
NC_017585	S. cattleya NRRL 8057	0.702
NC_011879 NZ_CM001010	A. chlorophenolicus A6	0.648
NZ_CM001019 NZ_CM000914	S. clavuligerus ATCC 27064 S. clavuligerus ATCC 27064	0.642
Cyanobactéries	5. clavallycrus 11100 21004	0.042
Chroococcales		
NC_009927	A. marina MBIC11017	0.582
NC_009926	A. marina MBIC11017	0.578
Deinococcus-THERMUS		
NC_017805	D. aobiensis I-0	0.812
NC_008010	D. geothermalis DSM 11300	0.622
Thermales		
NC_017588	T. thermophilus JL-18	0.557
NC_006462	T. thermophilus HB8	0.505
Clostridiales		
NC_012654	C. botulinum Ba4 str. 657	0.531
NC_010418	C. botulinum A3 str. Loch Maree	0.531
Alphaprotéobactérie	S	
Rhizobiales		0.007
NC_017323	S. meliloti BL225C	0.961
NC 003078	5. meliloti 1021	0.960
NC 017326	S. meliloti SM11	0.947
NC_009620	S. medicae WSM419	0.942
NC_018683	S. meliloti Rm41	0.922
NC_016815	S. fredii HH103	0.915
NC_012586	S. fredii NGR234	0.894
NC_017327	S. meliloti SM11 S. meliloti PI 225C	0.877
NC 009621	S. medicae WSM419	0.830
NC_003037	S. meliloti 1021	0.818
NC_010997	R. etli CIAT 652	0.792
NC_011368 R.	leguminosarum bv. trifolii WSM2304	0.777
NC_012858 R.	leguminosarum bv. trifolii WSM1325	0.741
NC_008384	R. leguminosarum bv. viciae 3841	0.731
NC 008378	R. etti CFN 42 R. leauminosarum by viciae 3841	0.725
NC_012848 R.	leguminosarum by, trifolii WSM1325	0.711
NC_010998	R. etli CIAT 652	0.701
NC_011366 R.	leguminosarum bv. trifolii WSM2304	0.630
NC_015184	A. sp. H13-3	0.565
NC_007766	R. etli CFN 42	0.555
Rhodobactóralos	M. extorquens AM1	0.538
NC_008688	P. denitrificans PD1222	0.769
NC_008043	<i>R. sp.</i> TM1040	0.667
${f R}hodospirillales$	-	
NC_016594	A. brasilense Sp245	0.878
NC_017958	T. mobilis KA081020-065	0.797
NC 016585	A. sp. D010 A linoferum AB	0.732
NC_016587	A. lipoferum 4B	0.645
NC_016586	A. lipoferum 4B	0.609
NC_016595	A. brasilense Sp245	0.603
NC_016618	A. brasilense $Sp245$	0.591
NC_017966	T. mobilis KA081020-065	0.578
NC_013858	A. $sp.$ B510 A $sp.$ B510	0.545
Sphingomonadales	A. sp. D010	0.000
NC_015583	N. sp. PP1Y	0.523
Bêtaprotéobactéries		
Burkholderiales		
NC_007974	C. metallidurans CH34	0.883
NC 003206	к. solanacearum Po82 В solanacearum CMI1000	0.865
NC_016626	B. sp. YI23	0.846
NC_014310	R. solanacearum PSI07	0.827
NC_010625	B. phymatum STM815	0.733
NC_018696	B. phenoliruptrix BR3459a	0.663
NC_015727	C. necator N-1	0.575
NC_007336	K. eutropha JMP134	0.513
Entérobacteriales	E-0	
NC_014838	Pantoea sp. At-9b	0.527
	-	

(1	3) $\{E_{BECE}, E_{plasmide}\}^{it}$	
Actinobactéries	, (10202), pracinitacij	
Actinomycetal	es	
NC_016113	S. cattleya NRRL 8057	0.827
NC_017585	S. cattleya NRRL 8057	0.811
NZ_CM001019	S. clavuligerus ATCC 27064	0.716
NZ_CM000914	S. clavuligerus ATCC 27064	0.716
NC_011879	A. chlorophenolicus A6	0.671
NC 003903	S coelicolor A3(2)	0.552
NC 008269	B iostii BHA1	0.543
CVANODACTÉDIES	n. jostu mini	0.040
CYANOBACTERIES		
NC 000027	A maring MPIC11017	0.006
NC_009927	A. marina MBICI1017	0.906
NC-009926	A. marina MBIC11017	0.854
NC_009928	A. marina MBIC11017	0.742
NC_011737	C. sp. PCC 7424	0.610
NC_010474	S. sp. PCC 7002	0.602
NC_011738	C. sp. PCC 7424	0.550
NC_014534	C. sp. PCC 7822	0.530
Nostocales		
NC 010632	N nunctiforme PCC 73102	0 758
NC 007412	A suggistic ATCC 20412	0.158
DEPENDING	A. Variabilis AICC 29413	0.058
DEFERRIBACTERES		
Deferribacteral	es Di la canti	0.010
NC_013940	D. desulfuricans SSM1	0.618
Deinococcus-Thei	RMUS	
Deinococcales		
NC_017805	D. gobiensis I-0	0.901
NC_008010	D. geothermalis DSM 11300	0.853
NC_015169	D. proteolyticus MRP	0.751
NC_012528	D. deserti VCD115	0.703
NC 012529	D deserti VCD115	0.654
NC 017791	D. achiensis I.O.	0.619
NC 017771	D. sobiensis I-0	0.010
NC 010707	D. gootensis 1-0	0.008
NC_012527	D. aeserti VCD115	0.607
NC_000958	D. radiodurans R1	0.589
Thermales		
NC_{017588}	T. thermophilus JL-18	0.756
NC_017273	T. thermophilus SG0.5JP17-16	0.734
NC_006462	T. thermophilus HB8	0.606
NC 019387	T. oshimai JL-2	0.564
FIRMICUTES		
Bacillalos		
NC 011990	B carrens U2001 07	0.610
NC 010001	D. CETEUS FI3001.97	0.019
NC_010921	B. cereus	0.594
INC_010916	B. cereus	0.594
NC_011777	B. cereus AH820	0.594
NC_010180	B. weihenstephanensis KBAB4	0.580
NC_018689	B. thuringiensis MC28	0.557
NC_018688	B. thuringiensis MC28	0.55
NC 011775	B. cereus G9842	0.512
Clostridiales	D. 00/040 00012	0.012
NC 012780	E eligens ATCC 27750	0 711
NC 014994	D -11 7	0.711
NC_014824	n. uious (0.625
NC_012654	C. botulinum Ba4 str. 657	0.604
NC_010418	C. botulinum A3 str. Loch Maree	0.590
NC_014390	B. proteoclasticus B316	0.566
NC_012219	C. botulinum	0.515
NC_012946	C. botulinum D str. 1873	0.515
Thermomicrobia		
Thermomicrob	iales	
NC_011961	T. roseum DSM 5159	0.529
		0.010

(C) $\{E_{RECE}, E_{plasmide}\}^{it}$ (suite)	
Alphaprotéob	ACTÉRIES	
Caulobactér	ales	
NC_010335	C. sp. K31	0.506
NC 018701	S malilati Pm41	0.055
NC 017323	S. meliloti BL225C	0.952
NC 017326	S. meliloti SM11	0.947
NC_009620	S. medicae WSM419	0.946
NC_003078	S. meliloti 1021	0.934
NC_018683	S. meliloti Rm41	0.893
NC_012586	S. fredii NGR234	0.884
NC_016815	S. fredii HH103	0.877
NC_017324	S. meliloti BL225C	0.862
NC_009621	S. medicae WSM419	0.859
NC_017327	S. meluloti SM11	0.833
NC 011368	B leauminosarum by trifolii WSM2304	0.780
NC 012811	M. extorauens AM1	0.743
NC_012858	R. leauminosarum by, trifolii WSM1325	0.715
NC_010997	R. etli CIAT 652	0.711
NC_008384	R. leguminosarum bv. viciae 3841	0.693
NC_007765	R. etli CFN 42	0.692
NC_010998	R. etli CIAT 652	0.678
NC_012848	R. leguminosarum bv. trifolii WSM1325	0.675
NC_008378	R. leguminosarum bv. viciae 3841	0.647
NC_011366	R. leguminosarum bv. trifolii WSM2304	0.634
RC_007766	R. etti CFN 42	0.550
NC 008688	P denitrificans PD1222	0.816
NC_008043	<i>B. sp.</i> TM1040	0.715
Rhodospirill	ales	
NC_016594	A. brasilense Sp245	0.910
NC_017958	T. mobilis KA081020-065	0.884
NC_016586	A. lipoferum 4B	0.850
NC_016585	A. lipoferum 4B	0.843
NC_013855	A. sp. B510	0.812
NC_016587	A. lipoferum 4B	0.809
NC_016505	A. sp. B510	0.791
NC 016596	A. brasilense Sp245	0.743
NC 016618	A brasilense Sp245	0.716
NC_013857	A. sp. B510	0.680
NC_017957	T. mobilis KA081020-065	0.643
NC_017966	T. mobilis KA081020-065	0.630
NC_016623	A. lipoferum 4B	0.591
NC_013856	A. sp. B510	0.582
Sphingomon	adales	
NC_015583	N. sp. PP1Y	0.622
NC_009507	S. wittichii RW I	0.616
	5. japonicum 01205	0.909
Burkholderi	ales	
NC_003296	R. solanacearum GMI1000	0.936
NC_014310	R. solanacearum PSI07	0.922
NC_007974	C. metallidurans CH34	0.919
NC_017575	R. solanacearum Po82	0.915
NC_018696	B. phenoliruptrix BR3459a	0.833
NC_016626	B. sp. YI23	0.805
NC_010625	B. phymatum STM815	0.748
NC_015727	C. necator N-1	0.724
NC-010627	B. phymatum STM815	0.661
NC 010520	R. eutropha JMF134	0.584
NC 014120	B_{sp} CCCE1002	0.504
GAMMAPROTÉOR	BACTÉRIES	0.010
Entérobacte	riales	
NC_015062	R. sp. Y9602	0.688
NC_017060	R. aquatilis HX2	0.631
NC_014838	<i>P. sp.</i> At-9b	0.545
Acidithiobac	llales	
NC_015851	A. caldus SM-1	0.655

Classification supervisée

148

TABLE 7.4: Plasmides classés comme RECE par la procédure de classification $f_{ERT}^{E_{training}}(V_f^{R^{\{plasmide, RECE\}}})$ avec $E_{training} = \{E_{RECE}, E_{plasmide}\}$ (7.4a) et $E_{training} = \{E_{RECE}, E_{plasmide}\}^{it}$ (7.4b et 7.4c).

 $\label{eq:première} \begin{array}{l} \mbox{Première colonne : numéro d'accession $RefSeq$ du réplicon, deuxième colonne : espèce-hôte du réplicon, dernière colonne : probabilité d'appartenance à la classe "RECE". \end{array}$

7.6 Résultats et discussion

Le classifieur ERT a été choisi pour ses performances en comparaison aux classifieurs SVM, Naive et AD, ainsi que pour sa capacité à être plus généralisable que RF [Geurts et al., 2006]. Ce classifieur a été utilisé pour les procédures de classification des plasmides, chromosomes et génomes décrites précédemment.

• Les résultats élevés obtenus pour les CV_{score} et les OOB_{score} des différents classifieurs montrent l'efficacité des STIG dans la discrimination des réplicons selon leur type (Table 7.3). Les scores très élevés obtenus pour la classification chromosome/plasmide des réplicons souligne les différences fonctionnelles de la répartition des STIG chez ces deux types d'élément. La discrimination RECE/plasmide apparaît plus ambiguë en présentant des scores plus faibles (Table 7.3). Cependant, de façon générale, les probabilités élevées obtenues en moyenne dans la classification des RECE dans leur classe respective atteste de la pertinence d'utiliser les STIG pour la différentiation de ces éléments génomiques.

• Ces analyses ont permis d'identifier de potentiels RECE parmi les plasmides (Table 7.4). Le détail de ces RECE est discuté plus loin (§7.7). Néanmoins, compte tenu des nombreux éléments de la littérature existante suggérant qu'une part importante de ces réplicons présente des caractèristiques de "RECE", la pertinence de ces procédures de classification est confirmée.

• Les STIG ne permettent de discriminer que partiellement certains RECE. Pour les RECE de certains lignées (*Leptospira* notamment), les STIG utilisés ne permettent pas d'identifier clairement ces réplicons comme des RECE, une partie de ceux-ci pouvant se classer préférentiellement dans la classe "plasmide" et/ou avoir une faible probabilité d'appartenance à la classe "RECE" (Table 7.5).

Paracoccus	0.96	Chlora cido bacterium	0.88
Ochrobactrum	0.96	Ilyobacter	0.88
Ralstonia	0.95	Sphaerobacter	0.88
A stic cacaulis	0.95	A liivibrio	0.87
Photobacterium	0.95	Cyanothece	0.86
Cupriavidus	0.94	Butyrivibrio	0.83
Anabaena	0.94	Variovorax	0.83
Prevotella	0.92	Deinococcus	0.78
Brucella	0.92	Thermobaculum	0.78
Pseudoal teromonas	0.91	Vibrio	0.76
No cardiops is	0.90	Sphingobium	0.73
Sin or hizobium	0.90	Rhodobacter	0.69
A grobacterium	0.90	Leptospira	0.54
Burkholderia	0.89		

TABLE 7.5: Moyenne, par genre bactérien, des probabilités des RECE d'appartenir à la classe "RECE" obtenues par la procédure de classification : $f_{ERT}^{E_{training}}(V_f^{R^{\{RECE\}}})$ avec $E_{training} = \{E_{RECE}, E_{plasmide}\}^{it}$

Il est possible que les gènes liés aux STIG de ces génomes sont présent dans l'analyse en nombre insuffisant, cas des génomes de *Leptospira*, par exemple. On peut néanmoins souligner que, malgré le faible nombre d'attributs (six, *cf.* §5.3.2.1) décrivant les RECE de *Leptospira*, certains attributs, annotés ParA et ParB chromosomiques notamment, sont relativement caractéristiques d'un état non-plasmidique. Pour les RECE des *Vibrionaceae*, les protéines RtcB, spécifiques des RECE de cette famille, n'ont pas été prises en compte du fait de leur stricte spécificité envers cette famille bactérienne. Certaines structures spécifiques liées aux STIG non prises en compte (position et caractéristiques structurelles d'*ori* par exemple) pourraient être les clés pour permettre la discrimination de ces réplicons, (*cf.* §2.5).

• Une très faible part (huit) des réplicons extra-chromosomiques se classent comme "chromosome" (Table 7.6). Parmi eux, se retouvent les RECE de *Astic-cacaulis, Paracoccus* et *Prevotella*, confirmant les biais marqués observés pour la distribution des gènes des STIG de ces réplicons.

Néanmoins, seul le RECE de *P. intermedia* présente une probabilité très importante (P > 0.98) qui peut être expliquée par la structure particulière du génome de *P. intermedia* 17 (cf. §8.6). La présence du RECE de *N. dassonvillei* dans le groupe "chromosome" est très peu significative (P < 0.54). La présence, avec une faible probabilité, de mégaplasmides de *Methylbacterium extorquens* et *Azospirillum* est due au fait qu'ils comportent des gènes codant pour des protéines, annotées DnaG, DnaB, ParC ou ParE entre autres, qui sont très inhabituelles pour des réplicons extrachromosomiques (Table 6.6). Ces réplicons sont de plus identifiés comme de potentiels "RECE" avec une forte probabilité pour *Azospirillum* mais une probabilité non significative pour *M. extorquens*. Enfin, à l'exception du chromosome de *P. intermedia* 17, aucun chromosome n'est classé dans la classe "plasmide" (résultats non montrés).

TABLE 7.6: R	éplicons	extra-chromo	somiques	classés	comme	chromosome	par	la
procédure	$f_{ERT}^{E_{trainin}}$	${}^{g}(V_{f}^{R^{\{RECE,pla}})$	$^{smide})$ ave	ec E_{train}	$ing = \{E$	$E_{chr}, E_{plasmide}$	}.	

Actinobactéries			
Actinomycétales			
NC_014211	N. dassonvillei subsp. dassonvillei DSM 43111	RECE	0.539
Alphaprotéobactéries			
Caulobactérales			
NC_014817	A. excentricus CB 48	RECE	0.637
Rhizobiales			
NC_012811	$M \ extorquens \ AM1$	plasmide	0.669
Rhodobactérales			
NC_008687	P. denitrificans PD1222	RECE	0.778
Rhodospirillales			
NC_016594	A. brasilense Sp245	plasmide	0.774
D			
BACTEROIDETES			
Bacteroidales		DECE	0.004
NC_017861	P. intermedia 17	RECE	0.984
NC_014371	P. metaninogenica ATUU 25845	RECE	0.698
Cyanobactéries			
Nostocales			
NC_019439	Anabaena sp. 90	RECE	0.638

• Les attributs les plus discriminants dans la classification RECE/plasmide (Table 7.7) sont similaires à ceux identifiés comme "significatifs" dans l'analyse par régression logistique (Table 6.6). Des résultats comparables sont de plus observés pour les classifications obtenues avec les autres training-sets (résultats non montrés). Ces résultats confirment le pouvoir discriminant de certaines fonctions des STIG (cf. §6.4.3) dans la séparation des réplicons selon leur type.

TABLE 7.7 :	Importance des	s attributs	fonctionnels	s dans la	procédure	de classification
	$f_{ERT}^{E_{training}}$	avec E_{trai}	$i_{ning} = \{E_{RI}\}$	ECE, E_{pl}	$asmide$ it .	

Fonction	Score	Fonction	Score ⁻	Fonction	Score ^a
kegg_ftsE	0.1016	aclame_DNAhelicase	0.0048	aclame_PSK_ccd	0.0004
kegg_acrA	0.0733	kegg_dps	0.0046	kegg_mukF	0.0004
kegg_parA_soj	0.0696	kegg_ftsA	0.0044	kegg_smc	0.0003
kegg_hupB	0.0556	kegg_ssb	0.0042	aclame_Fis	0.0003
kegg_lrp	0.0529	kegg_parC	0.0038	kegg_ftsQ	0.0002
kegg_rob	0.0492	aclame_XerD	0.0037	aclame_PSK_HOK/SOK	0.0002
kegg_minD	0.0478	kegg_parE	0.0034	kegg_ftsB	0.0002
kegg_iciA	0.0424	kegg_ihfB_himD	0.0034	kegg_mreC	0.0002
kegg_ftsI	0.0392	kegg_hns	0.0033	aclame_Rop	0.0002
kegg_xerC	0.0258	kegg_dnaC	0.0032	kegg_mreD	0.0001
aclame_ATPase/tyrK/exoP	0.0243	kegg_dam	0.0032	aclame_DNAbinding	0.0001
kegg_cbpA	0.0238	kegg_ftsK_spoIIIE	0.0031	aclame_PSK_yacA	0.0001
kegg_xerD	0.0227	kegg_rodA_mrdB	0.0028	kegg_hda	0.0001
kegg_mrp	0.0223	aclame_PSK_epsilon-zeta	0.0027	kegg_gidA_mnmG_MTO1	0.0001
kegg_mreB	0.0209	kegg_sulA	0.0026	kegg_trmFO_gid	0.0001
kegg_parB_spo0J	0.0175	aclame_PSK_relBE	0.0025	aclame_cdsD	0.0001
aclame_RuvB	0.0162	aclame_PSK_phD-doc	0.0025	aclame_ParR_ParB	0.0
kegg_E3.5.1.28B_amiA_amiB_amiC	0.0151	kegg_hfq	0.0023	aclame_Rep	0.0
aclame_ParB	0.0145	aclame_RepA	0.0023	aclame_helicase	0.0
kegg_minC	0.0139	aclame_RepA_E_B	0.0023	aclame_RepR_S_E	0.0
kegg_minE	0.0137	kegg_dnaG	0.0019	kegg_divIVA	0.0
kegg_ftsX	0.0132	kegg_slmA_ttk	0.0019	kegg_racA	0.0
aclame_FtsK/SpoIIIE	0.0105	kegg_ftsZ	0.0019	kegg_dnaI	0.0
aclame_XerTyrosine	0.0102	kegg_scpA	0.0019	kegg_ezrA	0.0
aclame_Helicase	0.0101	kegg_dnaA	0.0017	kegg_hupA	0.0
aclame_PSK_HicAB	0.0099	aclame_CopG	0.0017	kegg_stpA	0.0
kegg_ftsW_spoVE	0.0097	aclame_TrfA	0.0016	aclame_PSK_parC	0.0
aclame_PSK_parDE	0.0094	aclame_PSK_mazEF	0.0016	kegg_ftsN	0.0
aclame_DNArepair	0.0083	aclame_primase_LtrC	0.0011	aclame_RepC_J_E	0.0
aclame_ParA/ParM	0.0083	kegg_tus_tau	0.0011	kegg_sepF	0.0
aclame_serinerecombinase	0.0082	aclame_DnaB	0.0010	kegg_diaA	0.0
kegg_fic	0.0080	kegg_gidB_rsmG	0.0008	aclame_plasmidmaintenance_PSK	0.0
aclame_PLdimerresolution	0.0077	kegg_zapA	0.0007	kegg_dnaB2_dnaB	0.0
kegg_dnaB	0.0071	kegg_fis	0.0007	kegg_seqA	0.0
aclame_RepC	0.0064	kegg_mukE	0.0006	kegg_ftsL	0.0
aclame_PSK_vapBC/vag	0.0058	aclame_RepA_BCopB	0.0005	kegg_divIC_divA	0.0
aclame_TyrosinerecOrfA	0.0052	kegg_mukB	0.0005	aclame_PSK_vapXD	0.0
kegg_scpB	0.0051	aclame_RepB	0.0005	kegg_zipA	0.0
aclame_PSK_higBA	0.0048	kegg_ihfA_himA	0.0004	aclame_RNApolymerase	0.0

^a Score de l'importance des attributs selon les variations de l'OOB_{score} (décrit dans [Breiman, 2001]).

• Choisies en tant qu'attributs des réplicons, les annotations fonctionnelles des protéines des STIG permettent une meilleure discrimination des types que les clusters des protéines STIG seuls. Les OOB_{score} (Tables 7.3 et 7.8) montrent une plus grande efficacité ($p_{value} \ll 0.05$ pour les tests de Kolmogorov-Smirnov) des attributs fonctionnels dans les classifications RECE/plasmide et chromosome/RECE, mais pas de différence significative pour la classification plasmide/chromosome.

TABLE 7.8: OOB_{score} obtenus avec la procédure de classification $f_{ERT}^{E_{training}}$ en utilisant les clusters protéiques comme attributs des réplicons.

 $\{ V^{R^{\{RECE\}}}, V^{R^{\{plasmide\}}} \}^{it} \text{ et } \{ V^{R^{\{RECE\}}}, V^{R^{\{chr\}}} \}^{it} \text{ désignent des procédures itératives d'échantillonnage aléatoire similaires à celles décrites §7.3. }$

${ m E_{training}}$	OOB_{score}	$\sigma_{OOB_{score}}$
$\{V^{R^{\{RECE\}}}, V^{R^{\{plasmide\}}}\}$	0,89	-
$\{V^{R^{\{RECE\}}}, V^{R^{\{plasmide\}}}\}^{it}$	$0,\!82$	0,02
$\{V^{R^{\{chr\}}}, V^{R^{\{plasmide\}}}\}$	$1,\!0$	-
$\{V^{R^{\{RECE\}}}, V^{R^{\{chr\}}}\}^{it}$	$0,\!95$	0,02

Ces résultats laissent présager que, par rapport aux STIG, les RECE se différencient des chromosomes et des plasmides plus par les spécificités fonctionnelles de leurs STIG, que structurellement, selon les homologies des séquences protéiques des STIG, témoignant des origines plasmidiques et/ou chromosomiques des RECE.

• Les classifieurs utilisés ainsi que les attributs sélectionnés ne permettent pas de discriminer significativement les génomes multipartites des génomes monopartites (Tables 7.3 et 7.9), rejoignant ainsi les conclusions présentées §6.4.3 où seulement de faibles biais sont detectés entre $K_{monopartite}$ et $K_{multipartite}$.

TABLE 7.9: Moyenne, par genre bactérien, des probabilités des différents génomes multipartites d'appartenir à la classe "multipartite", obtenues par la procédure de classification : $f_{ERT}^{E_{training}}(V_f^{G^{\{multipartite\}}})$ avec $E_{training} = \{G_{multipartite}, E_{monopartite}\}^{it}$.

Ochrobactrum	0.8690	Sphingobium	0.7355
Sin or hizobium	0.8610	Thermobaculum	0.722
Paracoccus	0.8550	Candidatus	0.7130
Photobacterium	0.8230	Deinococcus	0.7040
Variovorax	0.8110	Ilyobacter	0.6880
Cyanothece	0.8010	No cardiops is	0.6650
Anabaena	0.7970	Vibrio	0.6466
A liivibrio	0.7960	Brucella	0.6425
A grobacterium	0.7818	Rhodobacter	0.6317
A stic cacaulis	0.7730	Burkholderia	0.6271
Ralstonia	0.7540	Pseudoal teromonas	0.5925
Butyrivibrio	0.7540	Prevotella	0.5310
Cupriavidus	0.74650	Leptospira	0.4844
Sphaerobacter	0.7440		

Une majorité de génomes multipartites présente une faible probabilité (P < 0.70) d'être classée en tant que "multipartite" (Table 7.9). Cette constatation est encore plus frappante lors de l'utilisation de { $E_{monopartite}, E_{multipartite}$ } comme training set (et non de la procédure itérative (résultats non montrés)), où seulement un quart des génomes multipartites sont classés comme "multipartite". Cependant, à une exception près : *Leptospira*, l'ensemble des génomes multipartites est correctement classé (P > 0.5, Table 7.9), avec des probabilités élevées pour certains genres. Si ces observations ne découlent pas d'artefacts, *e.g.*, nombre réduit et faible diversité des génomes multipartites, attributs manquants, corrélations non prises en compte..., (*cf.* §6.4.4), alors il **existe des caractéristiques identifiables chez les STIG des génomes multipartites**. Les attributs fonctionnels pertinents dans la classification sont similaires (résultats non montrés) à ceux précédemment identifiés (Table 6.6) et englobent principalement les protéines de partition, de résolution de dimères (XerC, XerD) et certains régulateurs (Lrp, IciA...).

7.7 Les nouveaux RECE

Parmi les réplicons nouvellement identifiés comme RECE (Tables 7.4a et 7.4b), il existe dans la littérature des indices les rapprochant des RECE :

- Parmi les réplicons extra-chromosomiques d'Azospirillum, il a été suggéré qu'un des plasmides de A. brasilense est en fait essentiel [Acosta-Cruz et al., 2012]. Ce RECE est identifié par l'analyse ainsi que certains plasmides additionnels de A. brasilense, A. lipoferum et A. sp. B510. On peut faire ainsi les hypothèses que ces réplicons identifiés sont i) les homologues du RECE déjà identifié chez A. brasilense ou ii) des RECE "en devenir" (cf. §8.2.4).
- Il a été proposé récemment que différents réplicons extra-chromosomiques de *Rhizo-bium* (p42e, pA, pRL11, pRLG202 et pR132502) soient désignés par le terme de "chromosome secondaire" [Landeta et al., 2011; Villaseñor et al., 2011]. La nature essentielle de ces réplicons (NC_007765, NC_010998, NC_008384, NC_011366, NC_012858, respectivement) est confirmée par notre l'analyse.
- pSymA (Sinorhizobium spp.), comporte de nombreux gènes importants pour la fitness de l'organisme [Blanca-Ordóñez et al., 2010] et, bien que considéré comme nonessentiel, contribue grandement à la fitness de l'hôte [Galardini et al., 2013]. Dans le jeux de données initial, seul le génome de Sinorhizobium meliloti AK83 est annoté comme étant multipartite. Notre analyse permet d'identifier les réplicons extra-chromosomiques similaires aux deux RECE pSymA et pSymB chez les autres souches de S. meliloti ainsi que chez S. fredii et S. medicae.
- Les trois mégaplasmides, pTM1, pTM2 et pTM3 (> 600kb), sur les quatre réplicons extra-chromosomiques que le génome de *Tistrella mobilis* comporte sont identifiés comme des RECE par notre analyse. pTM2 et pTM3 possèdent des opérons ARNt et ARNr, soulignant leur importance dans le génome.
- Chez Butyrivibrio proteoclasticus, le plasmide pCY186 (NC_014390) est identifié (avec une faible probabilité : P = 0.56) comme un potentiel RECE. Ce réplicon comporte de nombreuses protéines impliquées dans la réplication chromosomique [Yeoman et al., 2011]. L'autre plasmide, pCY360 (NC_014389), également reconnu comme essentiel par Yeoman et al. [2011], n'est pas détecté comme un RECE dans notre analyse (mais est légèrement biaisé : P = 0.32).
- Les différents plasmides identifiés chez les *Burkholderiales* témoignent de la capacité des réplicons extra-chromosomiques de cette famille à s'échanger du matériel génétique [Maida et al., 2014] et à former des RECE à partir des plasmides [Passot et al., 2012]. Les réplicons de *Ralstonia solanacearum* (annoté monopartite dans RefSeq) identifiés comme RECE sont ainsi vraisemblablement les homologues des RECE des espèces multipartites de *Ralstonia*.
- Chez les Actinobactéries, certains génomes possèdent des mégaplasmides singuliers par leur linéarité et leur taille (≈ 1.5 Mb). Le mégaplasmide linéaire de Streptomyces cattleya (1.8 Mb), identifié comme RECE avec des scores important (P > 0.7), ainsi que celui de S. coelicolor, possèdent des gènes impliqués dans les voies de synthèse de divers antibiotiques et métabolites secondaires [Barbe et al., 2011; O'Rourke et al., 2009]. Le mégaplasmide linéaire de S. clavuligerus, également

RECE potentiel, est un vaste réservoir de gènes de voies métaboliques en contenant plus de 20% des gènes du génome et possèdant de nombreuses régulations croisées avec le chromosome [Medema et al., 2010]. Le chromosome de *S. clavuligerus* dépend de plus du gène *tap*, impliqué dans la réplication du télomère, codé par le plasmide. Même si aucun gène codé par le mégaplasmide semble appartenir au génome-coeur [Medema et al., 2010], il est envisageable que le mégaplasmide contribue très fortement à la *fitness* de l'organisme.

- Plusieurs plasmides de Acaryochloris marina (2 ou 3 selon l'analyse : pREB1, pREB2 et pREB3, Table 7.4) ont été identifiés comme des RECE potentiels. Ces mégaplasmides (de taille comprise entre 273 et 354 kb) codent tous pour des protéines clé du métabolisme [Swingley et al., 2008], ce qui laisse envisager qu'ils contribuent à la fitness de l'organisme.
- pAQ7, le plus grand (186 kb) plasmide de *Synechococcus sp. PCC 7002*, identifié comme RECE par notre analyse, est présent dans le génome en même nombre de copies que le chromosome [Xu, 2010].
- pRAHAQ01, le mégaplasmide de Rahnella sp. Y9602, RECE potentiel, a un pourcentage en G+C (52.1%) très proche de celui du chromosome (52.4%) [Martinezet al., 2012]. Une constatation similaire est faite pour le plasmide de R. aquatilis HX2 aussi identifié comme un RECE.
- Le mégaplasmide (821kb) de *Ruegeria* comporte des opérons ARNr ainsi que des gènes uniques [Moran et al., 2007]. Il est notable de constater qu'à l'exception de ce mégaplasmide chez *Ruegeria*, aucun plasmide des proches *Roseobacter* n'a été identifié par l'analyse.
- Le mégaplasmide (1.2Mb) de *Methylobacterium extorquens* AM1, identifié comme RECE et comme chromosome (Table 7.6), possède une région synténique de 130kb avec le chromosome, une ploïdie de 1, ainsi que des opérons ARNt [Vuilleumier et al., 2009].

Ces différents éléments montrent la pertinence des analyses de classification et leur capacité à identifier des réplicons présentant une composition biaisée en gènes des STIG et potentiellement intégrés dans le cycle cellulaire. Cependant, il est évident que tous les réplicons identifiés ne sont pas forcément catégoriquement essentiels pour leur hôte, et que tous les réplicons de type RECE non annotés jusqu'à présent n'ont pas été identifiés (le plasmide pCY360 de *Butyrivibrio* en témoigne).

Cette analyse fait ressortir les réplicons présentant un biais significatif en gènes des STIG, ce qui laisse supposer **une prédisposition des réplicons identifiés à exister dans le génome dans un état de stabilité supérieur à celui des "vrais" plasmides**. Parmi les RECE identifiés, certains d'entre eux présentent une ploïdie similaire à celle du chromosome du génome et différente de celles des plasmides (*cf. Synechococcus, Methylobacterium...*), pouvant suggérer une coordination du RECE et du chromosome. De plus, les différents éléments de la littérature concernant les réplicons identifiés vont dans le sens que ces réplicons, par leur composition particulière en gènes des STIG, présentent un certain degré de stabilisation dans le génome leur offrant un potentiel de futur RECE, même si aucun des gènes de ces réplicons ne semble être strictement essentiel pour l'hôte.
Chapitre 8

Analyses de synténie des génomes mono- et multipartites

Les résultats des Chapitres 4, 5 et 6 mettent en évidence une spécificité marquée des STIG des RECE par comparaison à ceux des chromosomes et plasmides, des tendances caractéristiques des STIG des génomes multipartites par rapport aux génomes monopartites ainsi que des particularités de certains éléments génomiques (génomes réduits et génomes de bactéries associés aux plantes). Il semble de plus que diverses tendances existent au sein des RECE : certains d'entre eux sont significativement plus proches des chromosomes que des plasmides, et inversement, d'autres montrent une plus grande proximité des plasmides. Pour mieux comprendre l'organisation génomique, les mécanismes d'intégration, de régulation et de formation de ces éléments, des études complémentaires de génomique comparative sont conduites. L'organisation générale des génomes multipartites est analysée par l'analyse des synténies et l'étude descriptive des données génomiques disponibles dans la littérature.

Les différents résultats de la littérature semblent converger fortement vers l'hypothèse H2 d'émergence des RECE présentée §2.5 selon laquelle les RECE sont le résultat d'un enrichissement en gènes "chromosomiques" d'un plasmide pré-existant. Cependant, certains RECE (*Asticcacaulis, Paracoccus, Prevotella, Sphaerobacter*) semblent se distinguer par une proximité surprenante des chromosomes et, de ce fait, contrastent avec les hypothèses établies. Le modèle de l'hypothèse H2 ne s'applique donc pas pour ces réplicons. Alternativement, le processus d'intégration du RECE est tellement avancé qu'une origine plasmidique devient difficile à tracer.

8.1 Analyses de synténie des génomes multipartites

Pour mieux comprendre l'organisation générale de ces génomes multipartites, des études de synténie ont été conduites et concernent principalement des comparaisons entre RECE et chromosomes. La synténie, dans un contexte biologique, est définie comme la colocalisation de différents loci génétiques au sein d'un même chromosome (ou réplicon, par extension) [Renwick, 1971]. Une étude de synténie (ou synténie partagée) entre différents réplicons ou génomes fait référence à la co-localisation de loci entre ces réplicons ou génomes. Les loci utilisés sont les emplacements sur les réplicons de gènes orthologues. Les traces de synténie entre réplicons ou génomes témoignent alors d'homologies fortes entre ces éléments et d'origines évolutives communes [Landeta et al., 2011; Roberts et al., 2008]. Au sein des génomes bactériens, la conservation de l'ordre des gènes, en dehors des opérons, est très bruitée même pour des espèces très proches [Wolf et al., 2001]. Ainsi, le degré de synténie entre deux éléments génomiques bactériens est relié indirectement au degré de divergence évolutive, de THG, et de conservation et aux potentielles relations fonctionnelles entre les loci considérés. [Harrison, 2011].

8.1.1 Protocole analytique et outils utilisés

Les études de synténie ont été réalisées en utilisant l'outil SynMap de la plate-forme web CoGe (http://genomevolution.org/CoGe). Les algorithmes *blastn* [Camacho et al., 2009] ou Last [Kiełbasa et al., 2011] sont utilisés dans le pipeline analytique de Syn-Map afin d'effectuer une comparaison *All-vs-All* des différents gènes des deux séquences génomiques de l'analyse. Last est un algorithme similaire à ceux de la suite Blast mais significativement plus rapide et moins précis. Après différentes étapes de pré-processing, SynMap utilise ensuite l'algorithme DAGchainer [Haas et al., 2004] afin d'identifier les séries de paires de gènes colinéaires. Nous avons utilisé le paramétrage par défaut : un seuil d' e_{value} de 10⁻³ pour Last, un nombre minimum de 5 gènes pour qu'une zone soit considérée comme colinéaire, et une distance maximale tolérée de 20 gènes [Haas et al., 2004].

8.1.2 Démarche de l'étude

Les analyses de synténie ont été effectuées en comparant les génomes multipartites avec ceux d'espèces bactériennes les plus proches possibles sur le plan évolutif et dont le génome complet est disponible. Typiquement, les génomes multipartites sont comparés au génome monopartite de l'espèce la plus proche taxonomiquement, puis à des espèces plus éloignées d'un point de vue évolutif. Les résultats sont présentés sous forme de *dotplot* où abscisses et ordonnées représentent les génomes mono- et multipartites, et les marqueurs verts figurent les régions de synténie identifiées. La taille du génome étudié, en abscisse du *dotplot*, sert de référence et la longueur de l'ordonnée est proportionnelle à la taille des différents réplicons du génome comparé.

8.1.3 Indice synténique

Pour mesurer le taux de synténie partagée entre réplicons, un indice de comparaison est calculé à partir du nombre relatif de nucléotides inclus dans une région synténique entre chromosome et RECE des génomes multipartites et les réplicons d'un génome monopartite de référence. Soit un génome $G = \{r_1, ..., r_n\}$ avec r_i ses réplicons et $N(r_i)$ le nombre de nucléotides de r_i . L'annotation r_{chr} désigne un chromosome, $r_{plasmide}$ un plasmide et r_{RECE} un RECE. Soit $G_{ref} = \{r_1^{ref}, ..., r_n^{ref}\}$ un génome de référence comparé à G. Lors de la comparaison de G avec G_{ref} , on définit par $S_{nucl}(r_i, r_j^{ref})$ le nombre de nucléotides considérés comme faisant partie d'une région synténique entre r_i et r_j^{ref} pour $r_i \in G$. L'indice synténique $S_G(r_i, r_j^{ref})$ de $r_i \in G$ par rapport à r_j^{ref} est alors défini par :

$$S_G(r_i, r_j^{ref}) = \frac{\frac{S_{nucl}(r_i, r_j^{ref})}{N(r_i)}}{\frac{S_{nucl}(r_{chr}, r_j^{ref})}{N(r_{chr})}}.100$$
(8.1)

avec r_{chr} le chromosome de G.

Cet indice exprime sous forme de pourcentage le rapport entre le nombre de nucléotides des régions synténiques entre r_i et r_j^{ref} rapporté à la taille de r_i , et le nombre de nucléotides synténiques entre r_{chr} et r_j^{ref} rapporté à la taille de r_{chr} . Une valeur de 100% indique que le réplicon r_i de G a le même nombre de nucléotides synténiques par unité de taille avec r_j^{ref} que le chromosome r_{chr} de G. Lorsque $S_{nucl}(r_{chr}, r_j^{ref}) = 0$, alors $S_G(r_i, r_j^{ref})$ ne peut être calculé ($S_G(r_i, r_j^{ref}) = n.a.$ pour "non applicable"), l'objectif étant de comparer le taux de synténie du chromosome par rapport à celui de réplicons additionnels. Ce type de mesure est similaire à celle employée par Bavishi et al. [2010], qui compare la taille totale des "Local Colinear Blocks" entre réplicons, obtenus par le logiciel Mauve [Darling et al., 2004].

8.2 Génomes des Alphaprotéobactéries

8.2.1 Analyse des Rhodobactérales

Le génome multipartite de *Paracoccus denitrificans* PD1222 est comparé au génome monopartite de *Dinoroseobacter shibae* DFL 12 (Figure 8.1a; Table 8.1a) et au génome multipartite de *Rhodobacter sphaeroides* ATCC 17029 (Figure 8.1b; Tables 8.1b et 8.2b). Les espèces *P. denitrificans* et *R. sphaeroides* sont plus proches entre elles que chacune de *D. shibae. blastn* est utilisé.

Le génome multipartite de R. sphaeroides ATCC 17029 est aussi comparé au génome monopartite de D. shibae DFL 12 (Figure 8.2; Table 8.2a).



(A) Synténie de Paracoccus denitrificans PD1222 (ordonnée) vs. Dinoroseobacter shibae DFL 12 (abscisse).



(B) Synténie de Paracoccus denitrificans PD1222 (ordonnée) vs. Rhodobacter sphaeroides ATCC 17029 (abscisse).

FIGURE 8.1: Synténie de Paracoccus denitrificans PD1222 vs. Dinoroseobacter shibae DFL 12 (8.1a) et Rhodobacter sphaeroides ATCC 17029 (8.1b).

TABLE 8.1: Valeurs de l'indice synténique S_G (éq. 8.1) des réplicons de *Paracoccus* denitrificans PD1222 par rapport à ceux des génomes de *Dinoroseobacter shibae* DFL 12 (8.1a) et de *Rhodobacter sphaeroides* ATCC 17029 (8.1b).

(A) P. denitrificans PD1222 (G) vs. D. shibae DFL 12 (G_{ref}).

$G \diagdown G_{ref}$	r_{chr}^{ref}	$r_{pDSHI01}^{ref}$	$r_{pDSHI02}^{ref}$	$r_{pDSHI03}^{ref}$	$r_{pDSHI04}^{ref}$	$r_{pDSHI05}^{ref}$
r_{chr}	100	n.a.	100	n.a.	n.a.	n.a.
r_{RECE}	74	n.a.	n.a.	0	n.a.	n.a.
$r_{plasmide}$	12	n.a.	677	n.a.	n.a.	n.a.

(B) P. denitrificans PD1222 (G) vs. R. sphaeroides ATCC 17029 (G_{ref}).

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	$r_{plasmide}^{ref}$
r_{chr}	100	100	100
r_{RECE}	71	59	133
$r_{plasmide}$	20	40	0

Il existe une synténie plus importante entre le RECE de P. denitrificans PD1222 et les chromosomes de D. shibae DFL 12 (Table 8.1a) et de R. sphaeroides ATCC 17029 (Table 8.1b) qu'entre le RECE de R. sphaeroides ATCC 17029 et les chromosomes de P. denitrificans PD1222 (Table 8.2b) et de D. shibae DFL 12 (Table 8.2a). Une faible synténie est de plus présente entre les RECE des deux espèces à génome multipartite P. denitrificans PD1222 et R. sphaeroides ATCC 17029 (Tables 8.1b et 8.2b). On peut donc supposer que **les deux RECE sont issus d'événements évolutifs distincts ou évoluent plus vite** (Figure 8.1b). En comparant le génome de P. denitrificans PD1222 avec des génomes phylogéniquement plus distants : Maricaulis mari MCS10 (Rhodobactérales) et Caulobacter crescentus CB15 (Caulobactérales), des valeurs similaires sont observées malgré des signaux synténiques beaucoup plus faibles (résultats non montrés).

Il a été suggéré que certains plasmides des Rhodobactérales, dont deux de D. shibae DFL 12, sont en fait des "chromids" en raison de leur origine de réplication atypique [Petersen et al., 2013]. Nos précédentes analyses n'ont pas identifié de spécificités caractéristiques des RECE parmi la majorité des réplicons des genres des Roseobacter/Dinoroseobacter et ces analyses synténiques n'indiquent pas davantage de possibles transferts entre ces réplicons et les chromosomes (Tables 8.1a et 8.2a).



FIGURE 8.2: Synténie de *R. sphaeroides* ATCC 17029 (ordonnée) vs. Dinoroseobacter shibae DFL 12 (abscisse).

TABLE 8.2: Valeurs de l'indice synténique S_G (éq. 8.1) entre les réplicons de *Rhodobacter* sphaeroides ATCC 17029 et ceux des génomes (G_{ref}) de *Dinoroseobacter shibae* DFL 12 (8.2a) et de *Paracoccus denitrificans* PD1222 (8.2b).

(A) R. sphaeroides ATCC 17029 (G) vs. D. shibae DFL 12 (G_{ref}).

$G \diagdown G_{ref}$	r_{chr}^{ref}	$r_{pDSHI01}^{ref}$	$r_{pDSHI02}^{ref}$	$r_{pDSHI03}^{ref}$	$r_{pDSHI04}^{ref}$	$r_{pDSHI05}^{ref}$
r_{chr}	100	n.a.	100	n.a.	n.a.	n.a.
r_{RECE}	17	n.a.	0	n.a.	n.a.	n.a.
$r_{plasmide}$	0	n.a.	0	n.a.	n.a.	n.a.

(B) R. sphaeroides ATCC 17029 (G) vs. P. denitrificans PD1222 (G_{ref}).

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	$r_{plasmide}^{ref}$
r_{chr}	100	100	100
r_{RECE}	50	42	110
$r_{plasmide}$	26	49	0

Les bactéries du genre *Ruegeria*, appartenant également à la famille des Rhodobacteraceae, sont des espèces marines. Les espèces R. pomeyori et R. sp. TM1040 possèdent chacune un mégaplasmide. Parmi les plasmides des *Rhodobacteraceae*, seul le mégaplasmide de R. sp. TM1040, avec celui de *Paracoccus*, présente des spécificités de RECE (*cf.* Chapitre 7). Le génome de *Ruegeria* sp. TM1040 a donc été comparé à celui de D. shibae *DFL* 12 (Figure 8.3; Table 8.3). blastn est utilisé.



FIGURE 8.3: Synténie de *Ruegeria* sp. TM1040 (ordonnée) vs. Dinoroseobacter shibae DFL 12 (abscisse).

TABLE 8.3: Valeur de l'indice synténique S_G (éq. 8.1) entre les réplicons de *Ruegeria* sp. TM1040 (G) et ceux du génome de *Dinoroseobacter shibae* DFL 12 (G_{ref}).

$G \diagdown G_{ref}$	r_{chr}^{ref}	$r_{pDSHI01}^{ref}$	$r_{pDSHI02}^{ref}$	$r_{pDSHI03}^{ref}$	$r_{pDSHI04}^{ref}$	$r_{pDSHI05}^{ref}$
r_{chr}	100	n.a.	100	n.a.	n.a.	n.a.
$r_{megaplasmide}$	62	n.a.	0	n.a.	n.a.	n.a.
$r_{plasmide}$	23	n.a.	126	n.a.	n.a.	n.a.

Le mégaplasmide de Ruegeria sp. TM1040 possède de larges zones synténiques avec le chromosome de D. shibae DFL 12. La synténie est largement plus importante que dans le cas du RECE de R. sphaeroides avec le même chromosome de D. shibae relativement à la synténie partagée avec les chromosomes de Ruegeria sp. TM1040 et de R. sphaeroides (résultats non montrés). Cette caractéristique est de plus retrouvée en comparant le génome de Ruegeria sp. TM1040 à celui de P. denitrificans PD1222 (résultats non montrés). Le mégaplasmide de Ruegeria sp. TM1040 a donc une très forte probabilité d'être un RECE. Par contre, aucune région de synténie n'est detectée avec le RECE de P. denitrificans PD1222, ce qui suggère que le mégaplasmide/RECE de Ruegeria sp. TM1040 et le RECE de P. denitrificans PD1222 ont des origines différentes.

8.2.2 Analyse des Caulobactérales et des Sphingomonadales

8.2.2.1 Caulobactérales

Le génome d'Asticcacaulis excentricus CB 48 a été comparé à ceux de Caulobacter crescentus CB15, d'Erythrobacter litoralis HTCC2594 et de Zymomonas mobilis subsp. mobilis NCIMB 11163 (Figure 8.4). Les genres Asticaccaulis et Caulobacter appartiennent à la famille des Caulobacteraceae alors que Erythrobacter et Zymomonas sont des Sphingomonadaceae. blastn est utilisé.



(A) A. excentricus CB 48 (ordonnée) vs. Caulobacter crescentus CB15 (abscisse).



(B) A. excentricus CB 48 (ordonnée) vs. Zymomonas mobilis subsp. mobilis NCIMB 11163 (abscisse).



(C) A. excentricus CB 48 (ordonnée) vs. Erythrobacter litoralis HTCC2594 (abscisse).

FIGURE 8.4: Synténie entre Asticcacaulis excentricus CB 48 (abscisse) et Caulobacter crescentus CB15 (ordonnée; 8.4a), Zymomonas mobilis subsp. mobilis NCIMB 11163 (ordonnée; 8.4b), et Erythrobacter litoralis HTCC2594 (ordonnée; 8.4c).

TABLE 8.4: Valeurs de l'indice synténique S_G (éq. 8.1) entre les réplicons du génome d'Asticcacaulis excentricus CB 48 et ceux des génomes de Caulobacter crescentus CB15 (8.4a), Erythrobacter litoralis HTCC2594 (8.4b), et Zymomonas mobilis subsp. mobilis NCIMB 11163 (8.4c).

(A) A. excentricus CB 48 (G) vs. C. crescentus CB15 (G_{ref}).

(B) A. excentricus CB 48 (G) vs. E. litoralis HTCC2594 (G_{ref}).

$G \diagdown G_{ref}$	r_{chr}^{ref}	$G\diagdown G_{ref}$	r_{chr}^{ref}
r_{chr}	100	r_{chr}	100
r_{RECE}	170	r_{RECE}	194
$r_{pASTEX01}$	40	$r_{pASTEX01}$	0
$r_{pASTEX02}$	0	$r_{pASTEX02}$	0

(C) A. excentricus CB 48 (G) vs.Z. mobilis NCIMB 11163 (G_{ref}).

$G \diagdown G_{ref}$	r_{Mpl}^{ref}	$r_{pZA1001}^{ref}$	$r_{pZA1002}^{ref}$	$r_{pZA1003}^{ref}$
r_{chr}	100	n.a.	n.a.	n.a.
r_{RECE}	470	n.a.	n.a.	n.a.
$r_{pASTEX01}$	0	n.a.	n.a.	n.a.
$r_{pASTEX02}$	0	n.a.	n.a.	n.a.

De façon surprenante, les valeurs de l'indice S_G sont presque deux fois supérieures pour le RECE que pour le chromosome d'Asticcacaulis excentricus CB 48 (Figure 8.4) indiquant que la co-linéarité est quasiment deux fois plus conservée sur le RECE. La zone de synténie du RECE d'A. excentricus CB 48 avec le chromosome de C. crescentus CB15 semble se superposer à une zone sur le chromosome d'A. excentricus CB 48 où il y a une absence de synténie avec le chromosome d'A. excentricus CB 48 où il y a une absence de synténie avec le chromosome de C. crescentus CB15 (Figure 8.4a). Plus étonnant, on observe, comparativement au chromosome d'Asticcacaulis, un niveau de synténie similaire du RECE avec le chromosome d'Erythrobacter litoralis HTCC2594 (Table 8.5b) et une synténie 5 fois supérieure avec le chromosome de Zymomonas mobilis subsp. mobilis NCIMB 11163 (Table 8.4c). Cependant, compte-tenu des signaux synténiques très atténués pour les comparaisons des chromosomes avec Zymomonas et Erythrobacter, ces derniers résultats sont à interpréter avec prudence (cf. comparaison Caulobacter/Sphingobium ci-dessous).

8.2.2.2 Sphingomonadales

Le génome de Sphingobium japonicum NBRC 101211 est comparé à ceux de Caulobacter crescentus NA1000, Novosphingobium sp. PP1Y et Zymomonas mobilis subsp. pomaceae ATCC 29192. Sphingobium, Novosphingobium et Zymomonas appartiennent aux Sphingomonadaceae, Novosphingobium étant plus proche phylogéniquement de Sphingobium que Zymomonas. Caulobacter appartenant aux Caulobacteraceae est le plus distant de Sphingobium.



(A) Sphingobium japonicum NBRC 101211 (ordonnée) vs. Caulobacter crescentus NA1000 (abscisse).



(B) Sphingobium japonicum NBRC 101211 (ordonnée) vs. Novosphingobium sp. PP1Y (abscisse).



(C) Sphingobium japonicum NBRC 101211 (ordonnée) vs. Zymomonas mobilis subsp. pomaceae ATCC 29192 (abscisse).

FIGURE 8.5: Synténie entre Sphingobium japonicum NBRC 101211 (ordonnée) et Caulobacter crescentus NA1000 (abcisse; 8.5a), Novosphingobium sp. PP1Y (abcisse; 8.5b) et Zymomonas mobilis subsp. pomaceae ATCC 29192 (abcisse; 8.5c).

TABLE 8.5: Valeurs de l'indice synténique S_G (éq. 8.1) entre les réplicons du génome de Sphingobium japonicum NBRC 101211 et ceux des génomes de Caulobacter crescentus NA1000 (8.5a), Novosphingobium sp. PP1Y (8.5b) et Zymomonas mobilis subsp. pomaceae ATCC 29192 (8.5c).

(A) S. japonicum NBRC 101211 (G) vs. Caulobacter crescentus NA1000 (G_{ref}).

(B) S. japonicum NBRC 101211 (G) vs. Novosphingobium sp. PP1Y (G_{ref}).

$G \diagdown G_{ref}$	r_{chr}^{ref}	$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{Mpl}^{ref}	r_{Lpl}^{ref}	r_{Spl}^{ref}
r_{chr}	100	r_{chr}	100	100	100	100
r_{RECE}	96	r_{RECE}	70	0	980	0
r_{pUT1}	0	r_{pUT1}	0	0	0	0
r_{pUT2}	0	r_{pUT2}	0	0	0	0
r_{pCHQ1}	0	r_{pCHQ1}	249	1041	0	0

(C) S. japonicum NBRC 101211 (G) vs. Z. mobilis ATCC 29192 (G_{ref}).

$G \diagdown G_{ref}$	r_{Mpl}^{ref}	$r_{pZYMOP01}^{ref}$	$r_{pZYMOP02}^{ref}$
r_{chr}	100	n.a.	n.a.
r_{RECE}	17	n.a.	n.a.
r_{pUT1}	0	n.a.	n.a.
r_{pUT2}	0	n.a.	n.a.
r_{pCHQ1}	0	n.a.	n.a.

La synténie du RECE de Sphingobium, relativement importante avec le chromosome de Novosphingobium, devient faible avec le chromosome Zymomonas et disparaît complètement avec celui d'Erythrobacter litoralis HTCC2594, une autre Sphingomonadaceae (résultats non montrés). Elle reste cependant relativement importante avec le chromosome de *Caulobacter*, représentant d'une autre famille (Caulobacteraceae). Il apparaît clair qu'il existe une zone nucléotidique commune entre le RECE et le plasmide de Novosphingobium (Figure 8.5b), témoignant de l'origine plasmidique du RECE. De nombreuses zones synténiques sont de plus trouvées entre le chromosome de Novosphingobium et le plasmide pCHQ1 de Sphingobium, suggérant l'intégration d'un certain nombre de gènes plasmidiques dans le chromosome de Novosphingobium. Le plasmide pCHQ1 est de plus identifié, avec un faible indice de confiance, comme potentiel RECE (cf. Chapitre 7) indiquant une éventuelle transition plasmide/RECE pour ce réplicon. Il est assez intriguant de retrouver un indice synténique relativement fort entre le chromosome de *Caulobacter* et le RECE de *Sphingobium*. Cette synténie est due, entre autres, à l'identification de gènes d'ARN ribosomiques. Compte-tenu de la faible synténie entre les chromosomes de Sphingobium et Caulobacter, on peut considérer que cette valeur de l'indice synténique n'est pas forcément représentative d'une conservation de région synténique chez le RECE car aucune synténie n'est retrouvée avec les réplicons, RECE et chromosome, de Zymomonas et le chromosome de Erythrobacter, mais reflète plutôt la faible synténie entre les deux chromosomes.

8.2.3 Analyse des Rhizobiales

Le génome de Brucella melitensis biovar Abortus 2308 est comparé aux génomes de Bartonella australis Aust/NH1, Mesorhizobium loti MAFF303099, Sinorhizobium meliloti 1021, Rhizobium etli CFN 42, Agrobacterium vitis S4, Agrobacterium tumefaciens C58 et Bradyrhizobium japonicum USDA 110, bactéries qui font toutes partie des Rhizobiales (Figure 8.6). Last est utilisé en raison de la grande taille de certains génomes (Figure 8.7 et Table 8.6).

La plupart de ces espèces sont connues pour leur capacité à fixer le di-azote atmosphérique et à s'associer symbiotiquement avec les plantes de la famille des Légumineuses au niveau de leurs racines et parfois de leur tiges. Ces espèces sont de plus les hôtes de RECE, plasmides et mégaplasmides caractéristiques [Pinto et al., 2012], ayant une origine de réplication spécifique, comprenant un opéron repABC. L'histoire évolutive de leur génome apparaît étroitement liée à celle de leurs mégaplasmides. En particulier, il a été suggéré que les RECE des Rhizobiales dérivent d'un mégaplasmide repABC ancestral ("plasmide ITR" pour Intragenomic Translocation Recipient) ayant capturé certains gènes du chromosome [Slater et al., 2009] (Figure 8.6).



FIGURE 8.6: Hypothèse d'évolution du plasmide ITR chez les Rhizobiales. Plasmide ITR : transformé en RECE (bleu), intégré dans le chromosome (rouge), ou perdu (vert). Adaptée de [Slater et al., 2009].

Les plasmides/mégaplasmides repABC ne suivent cependant pas une évolution verticale et sont soumis à de nombreux réarrangements et échanges inter- et intragénomiques [Castillo-Ramírez et al., 2009; Slater et al., 2009]. De plus, de nombreuses études ont montré la plus grande variabilité des RECE des Rhizobiales par rapport aux chromosomes [Bavishi et al., 2010; Slater et al., 2009]. Brucella melitensis est un endosymbiote facultatif ayant un génome relativement petit (3.1 Mb), cette taille réduite pouvant être le résultat de leur état endosymbiotique [Wattam et al., 2009]. Les transferts des gènes chromosomiques sur le plasmide ITR semblent avoir eu lieu indépendamment chez les Brucellaceae : 25 clusters de gènes transférés du chromosome au RECE identifiés en comparaison, chez les Rhizobiaceae, d'un seul cluster commun identifié sur pSymB de S. meliloti et d'au moins deux clusters communs chez Rhizobium [Slater et al., 2009]. Rhizobium et Agrobacterium ont des plasmides ITR relativement proches, mais le RECE d'A. tumefaciens est linéaire contrairement à celui d'A. vitis [Lassalle, 2013]. Enfin, chez Bradyrhizobium et Mesorhizobium, le plasmide ITR serait intégré dans le chromosome, alors que pour d'autres espèces (dont Bartonella) le plasmide ITR aurait été perdu [Slater et al., 2009]. Les Rhizobiales représentent donc un jeu de données essentiel dans la comparaison des taux de synténie entre génomes multi- et monopartites.



(A) B. melitensis 2308 (ordonnée) vs. B. japonicum USDA 110 (abscisse).

(B) B. melitensis 2308 (ordonnée) vs. M. loti MAFF303099 (abscisse).



(C) B. melitensis 2308 (ordonnée) vs. S. meliloti 1021 (abscisse).

(D) *B. melitensis* 2308 (ordonnée) vs. *R. etli* CFN 42 (abscisse).



(E) B. melitensis 2308 (ordonnée) vs. B. australis Aust/NH1 (abscisse).



(F) B. melitensis 2308 (ordonnée) vs. A. tumefaciens (abscisse).

(G) B. melitensis 2308 (ordonnée) vs. A. vitis S4 (abscisse).

FIGURE 8.7: Synténie de Brucella melitensis bv. Abortus 2308 avec les génomes monopartites de Bradyrhizobium japonicum USDA 110 (8.7a), Mesorhizobium loti MAFF303099 (8.7b), Sinorhizobium meliloti 1021 (8.7c) et Rhizobium etli CFN 42 (8.7d), Bartonella australis Aust/NH1 (8.7e), et les génomes multipartites Agrobacterium tumefaciens C58 (8.7f) et Agrobacterium vitis S4 (8.7g).

TABLE 8.6: Valeurs de l'indice synténique S_G (éq. 8.1) pour les réplicons du génome *G* de *Brucella melitensis* bv. Abortus 2308 comparés aux génomes G_{ref} de *Bradyrhizobium japonicum* USDA 110 (8.6a). *Mesorhizobium loti* MAFF303099 (8.6b), *Bartonella australis* Aust/NH1 (8.6c), *Sinorhizobium meliloti* 1021 (8.6d), *Rhizobium etli* CFN 42 (8.6e), *Agrobacterium tumefaciens* C58 (8.6f) et *Agrobacterium vitis* S4 (8.6g).

(A) B. melitensis 2308 vs. B. japonicum USDA110

(B) B. melitensis 2308 vs. M. loti MAFF303099

$G \diagdown G_{ref}$	r_{chr}^{ref}	
r_{chr}	100	
r_{RECE}	42	

(D) B. melitensis 2308 vs. S. meliloti 1021

(C) B. melitensis 2308 vs. B. australis Aust/NH1

$G \diagdown G_{ref}$	r_{chr}^{ref}	$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{pSymA}^{ref}	r_{pSymB}^{ref}
r_{chr}	100	r_{chr}	100	100	100
r_{RECE}	21	r_{RECE}	44	173	161

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{p42a}^{ref}	r_{p42b}^{ref}	r_{p42c}^{ref}	r_{p42d}^{ref}	r_{p42e}^{ref}	r_{p42f}^{ref}
r_{chr}	100	n.a.	100	n.a.	100	100	100
r_{RECE}	40	n.a.	0	n.a.	0	788	238

(E) B. melitensis 2308 vs. R. etli CFN 42

(F) B. melitensis 2308 vs. A. tumefaciens C58

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}
r_{chr}	100	100
r_{RECE}	36	164

(G) B. melitensis 2308 vs. A. vitis S4

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	r_{pTiS4}^{ref}	r_{pAtS4a}^{ref}	r_{pAtS4b}^{ref}	r_{pAtS4c}^{ref}	r_{pAtS4e}^{ref}
r_{chr}	100	100	n.a.	n.a.	n.a.	n.a.	n.a.
r_{RECE}	38	465	n.a.	n.a.	n.a.	n.a.	n.a.

Les valeurs de l'indice synténique entre *Brucella* et *Bartonella* (Table 8.6c) sont proches de celles obtenues pour les RECE de *Rhodobacter* (Table 8.2) et des Burkholderiales (*cf.* Tables 8.8 et 8.9 ci-après), mais aussi pour les plasmides de *Paracoccus* (Table 8.1b) et de *Ruegeria* (Table 8.3). Si *Bartonella* a "perdu" le plasmide ITR, on peut considérer que la valeur de l'indice synténique obtenue entre le chromosome de *Bartonella* et le RECE de *Brucella* (Table 8.6) reflète un nombre d'échanges intragénomiques "classiques" entre chromosomes et réplicons extrachromosomiques (plasmides ou RECE).

Dans le cas de *Brucella*, ces échanges correspondent alors probablement aux gènes transférés sur le RECE et identifiés par Slater *et al.* [Slater et al., 2009].

Sous l'hypothèse que le plasmide ITR a été intégré dans les génomes ancestraux de

Bradyrhizobium et *Mesorhizobium*, on peut s'attendre à trouver des valeurs d'indice synténique plus élevées pour la comparaison du RECE de *Brucella* avec ces génomes. La confirmation de cette hypothèse pour les génomes de *M. loti* MAFF303099 (Table 8.6b) et *B. japonicum* USDA 110 (Table 8.6a), génomes phylogéniquement plus éloignés de *Brucella* que *Bartonella*, laisse penser que **des valeurs élevées de l'indice synténique témoignent de l'histoire évolutive commune de ces génomes avec le RECE de Brucella** (dérivant du plasmide ITR).

Ces tendances semblent se confirmer pour des génomes phylogéniquement plus ou moins distants (Figure 8.6). Les études synténiques entre . melitensis et Sinorhizobium (Figure 8.7c et Table 8.6d) montrent que les courtes régions synténiques partagées entre le RECE de Brucella et les plasmides pSymA et pSymB de Sinorhizobium sont deux fois plus importantes que l'étendue des synténies entre le chromosome de Brucella et les plasmides de Sinorhizobium, ce qui montre l'héritage nucléotidique commun entre RECE de B. melitensis et les plasmides pSymA et pSymB de S. meliloti. Une des explications possibles à la valeur élevée de l'indice synténique entre RECE de Brucella et plasmides ITR des Brucella sont différents de ceux transférés sur pSymB [Slater et al., 2009].

La comparaison avec les génomes de *Rhizobium* et *Agrobacterium* renforce cette tendance : les réplicons extrachromosomiques de ces espèces ont une valeur d'indice synténique avec le RECE de *Brucella* très supérieure à celles des comparaisons avec les chromosomes, témoignant ainsi de l'origine commune de ces réplicons extrachromosomiques à partir d'un plasmide ITR ancestral.

8.2.4 Analyse des Rhodospirillales

L'ordre des Rhodospirillales contient plusieurs espèces d'Azospirillum et de Tistrella qui hébergent des mégaplasmides de type RECE selon notre analyse (Chapitre 7, Table 7.4). Le génome d'Azospirillum brasilense Sp245 est comparé au génome de Rhodospirillum centenum ATCC 51521 (Figure 8.8). Tout comme de nombreuses Alphaprotéobactéries, Azospirillum brasilense fixe l'azote atmosphérique N_2 et est présent au niveau de la rhizosphère de certaines plantes. Rhodospirillum est également membre des Rhodospirillales. Les génomes des espèces du genre Azospirillum sont connus pour être multi-réplicons et, récemment, un des plasmides de Azospirillum brasilense a été classé parmi les "chromid" [Acosta-Cruz et al., 2012].

Une synténie importante existe entre le chromosome de *R. centenum* et le plus large plasmide d'*A. brasilense*, renforçant le caractère de RECE d'AZOBRp1 (NC_016594). Des synténies moins marquées existent aussi entre les plasmides additionnels et le chromosome indiquant de probables transferts latéraux. Ces résultats renforcent ceux présentés par Acosta-Cruz *et al.* soutenant fortement la présence de RECE d'origine plasmidique dans le genre *Azospirillum* [Acosta-Cruz et al., 2012]. On peut supposer que, comme pour les Burkholderiales [Passot et al., 2012], les divers réplicons extra-chromosomiques d'*Azospirillum* présentent des stades variables d'intégration dans le génome stable, observables par leur degré de synténie avec le génome de *Rhodospirillum centenum* ATCC 51521.



FIGURE 8.8: Synténie d'Azospirillum brasilense Sp245 (abscisse) vs. Rhodospirillum centenum ATCC 51521 (ordonnée).

TABLE 8.7: Valeurs de l'indice synténique S_G (éq. 8.1) pour les réplicons du génome *G* d'*Azospirillum brasilense* Sp245 par rapport au génome G_{ref} de *Rhodospirillum centenum* ATCC 51521.

$G \diagdown G_{ref}$	r_{chr}^{ref}
r_{chr}	100
$r_{AZOBRp1}$	63
$r_{AZOBRp2}$	23
$r_{AZOBRp3}$	9
$r_{AZOBRp4}$	19
$r_{AZOBRp5}$	0
$r_{AZOBRp6}$	0

8.3 Génomes des Bêtaprotéobactéries

La majorité des génomes multipartites référencés se trouve au sein des Burkholderiales, parmi les genres Burkholderia, Ralstonia et Cupriavidus. Les génomes de B. multivorans ATCC 17616 et de B. pseudomallei sont comparés aux génomes de Bordetella bronchiseptica 253, Ralstonia eutropha H16 et de Polynucleobacter necessarius STIR1 (Figure 8.9). De même, le génome de Ralstonia eutropha H16 est comparé à ceux de Bordetella bronchiseptica 253 et Polynucleobacter necessarius STIR1, ainsi qu'à celui de Janthinobacterium sp. Marseille, une espèce appartenant à l'ordre des Burkholderiales et à la famille des Oxalobacteraceae [Hornung et al., 2013]. Last est utilisé à cause de la grande taille des génomes comparés.

Les *Polynucleobacter* ont des génomes monopartites et sont membres, avec *Burkholderia*, des Burkholderiaceae. *Polynucleobacter necessarius* STIR1 est un endosymbiote d'*Euplotes aediculatus* et possède un génome de petite taille (1.5 Mb) [Hahn et al., 2009]. Les autres espèces à génomes multipartites parmi les Burkholderiales sont des *Ralsto-nia/Cupriavidus* de la famille des Ralstoniaceae, proche des Burkholderiaceae, et *Vario-vorax*, membre des Comamonadaceae. Les Ralstoniaceae comportent plusieurs espèces à génome multipartite ou dont le genome comprend des mégaplasmides structurellement proches des RECE [Passot et al., 2012]. *Variovorax* est actuellement le seul exemple de génome multipartite chez les Comamonadaceae. Les espèces du genre *Bordetella*, quant à elles, ont des génomes monopartites et font partie de la famille des Alcaligenaceae dans l'ordre des Burkholderiales.



(A) B. multivorans ATCC 17616 (abscisse) vs. B. bronchiseptica 253 (ordonnée).



(B) B. pseudomallei 668 (abscisse) vs. B. bronchiseptica 253 (ordonnée).



(C) B. multivorans ATCC 17616 (abscisse) vs. P. necessarius STIR1 (ordonnée).



(D) B. multivorans ATCC 17616 (abscisse) vs. R. eutropha H16 (ordonnée).

FIGURE 8.9: Synténie entre Burkholderia multivorans ATCC 17616 et Burkholderia pseudomallei 668 et Bordetella bronchiseptica 253 (8.9a et 8.9b, respectivement), et entre Burkholderia multivorans ATCC 17616 et Polynucleobacter necessarius STIR1 (8.9c) et Ralstonia eutropha H16 (8.9d). TABLE 8.8: Valeurs de l'indice synténique S_G (éq. 8.1) pour les réplicons du génome G de Burkholderia multivorans ATCC 17616 (8.8a) et de B. pseudomallei 668 (8.8b) par rapport au génome G_{ref} de Bordetella bronchiseptica 253, et pour les réplicons du génome G de B. multivorans ATCC 17616 par rapport aux génomes G_{ref} de Polynucleobacter necessarius STIR1 (8.8c) et Ralstonia eutropha H16 (8.8d).

(A) B. multivorans ATCC 17616 vs. B. bronchiseptica 253.

$G \diagdown G_{ref}$	r_{chr}^{ref}
r_{chr}	100
r_{RECE1}	21
r_{RECE2}	23
r_{pTGL1}	0
$r_{pBMUL01}$	0

(B) B. pseudomallei 668 vs. B. bronchiseptica 253.

$G \diagdown G_{ref}$	r_{chr}^{ref}
r_{chr}	100
r_{RECE}	19

(C)	$B.\ multivorans$	ATCC 1761	6 vs.	P.	necessarius
Ì,		STIR1.			

(D) B. multivorans ATCC 17616 vs. R. eutropha H16.

$G \diagdown G_{ref}$	r_{chr}^{ref}	$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	r_{pGH1}^{ref}
r_{chr}	100	r_{chr}	100	100	n.a.
r_{RECE1}	9	r_{RECE1}	15	122	n.a.
r_{RECE2}	3	r_{RECE2}	9	167	n.a.
r_{pTGL1}	0	r_{pTGL1}	21	0	n.a.
$r_{pBMUL01}$	0	$r_{pBMUL01}$	0	0	n.a.

Les espèces du genre Burkholderia ont quasiment toutes des génomes multipartites. Il existe à ce jour une unique exception : le génome de B. rhizoxinica. Selon les espèces de Burkholderia, un ou deux RECE ont été identifiés. Ces réplicons sont généralement considérés comme étant issus de l'adaptation de mégaplasmides originels (cf. Chapitre 2, §2.3 et §2.5). Les réplicons des génomes de Burkholderia possèdent un haut taux d'échange de gènes par TGH [Maida et al., 2014]. Les duplications géniques en résultant interviennent non seulement aux niveaux inter-chromosomique ou inter-plasmidique mais aussi entre chromosome et plasmide [Passot et al., 2012]. Ces réarrangements sont proposés comme constituant un mécanisme fondamental dans l'évolution des génomes [Maida et al., 2014]. Enfin, il semble que des échanges inter-génomes ont lieu avec des espèces proches phylogéniquement [Maida et al., 2014].

Il est intéressant de constater que, malgré la nature hautement plastique des réplicons de Burkholderia, seulement une faible synténie existe entre les RECE de Burkholderia et les chromosomes d'espèces proches phylogéniquement en comparaison de la synténie existant entre chromosomes (Table 8.8). L'absence de ressemblance RECE/chromosome est confirmée par la comparaison avec le génome de Polynucleobacter necessarius STIR1 (Figure 8.9c et Table 8.8c). P. necessarius possède un génome réduit de par son écologie endosymbiotique et son chromosome a donc tendance à n'abriter que des gènes de type chromosomique et essentiels.

Enfin, la comparaison de B. multivorans ATCC 17616 avec Ralstonia eutropha H16

(Figure 8.9d et Table 8.8d) montre une forte synténie entre les différents RECE, témoignant ainsi de leur origine commune. Les résultats mettent en évidence différents transferts entre RECE et chromosome spécifiques à *Ralstonia* ou à *Burkholderia*.

De façon surprenante, le RECE de *R. eutropha* H16 possède de larges zones synténiques avec le chromosome de *Bordetella bronchiseptica* 253 (Figure 8.10a et Table 8.9a), et par contre, très peu avec les génomes de *Janthinobacterium* sp. Marseille (Figure 8.10b et Table 8.9b) et *Polynucleobacter necessarius* STIR1 (Figure 8.10c et Table 8.10c), qui sont phylogéniquement plus proches de *Ralstonia*.



(A) R. eutropha H16 (abscisse) vs. B. bronchiseptica 253 (ordonnée).



(B) R. eutropha H16 (abscisse) vs. Janthinobacterium sp. Marseille (ordonnée).



(C) R. eutropha H16 (abscisse) vs. P. necessarius STIR1 (ordonnée).

FIGURE 8.10: Synténie du génome de Ralstonia eutropha H16 avec les génomes monopartites de Bordetella bronchiseptica 253 (8.10a), Janthinobacterium sp. strain Marseille (8.10b), et Polynucleobacter necessarius STIR1 (8.10c).

TABLE 8.9: Valeurs de l'indice synténique S_G (éq. 8.1) pour les différents réplicons du génome G de Ralstonia eutropha H16 par rapport aux génomes G_{ref} de Bordetella bronchiseptica 253 (8.9a), Janthinobacterium sp. Marseille (8.9b) et Polynucleobacter necessarius STIR1 (8.9c).

(A) R. eutropha H chiseptica	A) R. eutropha H16 vs. B. bron- chiseptica 253.		(B) R. eutropha H16 vs. Janthi- nobacterium sp. Marseille.			(C) R. eutropha H16 vs. P. ne cessarius STIR1.			
$G \diagdown G_{ref}$	r_{chr}^{ref}	$G \diagdown G_{ref}$	r^{ref}_{chr}	G^{\sim}	$\backslash G_{ref}$	r_{chr}^{ref}	_		
r_{chr}	100	r_{chr}	100		r_{chr}	100			
r_{RECE}	71	r_{RECE}	17	r_{1}	RECE	4			
r_{pHG1}	9	r_{pHG1}	0	r_{i}	pHG1	6			

Des résultats similaires sont obtenus en comparant les génomes d'espèces additionnelles de Bordetella et de Ralstonia/Cupriavidus (résultats non montrés). Le fait cependant qu'une faible synténie est partagée entre le RECE de R. eutropha H16 et les génomes monopartites de Janthinobacterium et Polynucleobacter indiquerait un événement génomique spécifique au genre Bordetella. En rapprochant la valeur de l'indice de synténie obtenue entre le RECE de R. eutropha strain H16 et le chromosome de B. bronchiseptica 253 à celles obtenues pour le génome de Mesorhizobium, on peut penser que le génome ancestral des Bordetella a intégré une partie des gènes du réplicon ancestral des RECE de Ralstonia/-Cupriavidus. Cette hypothèse est à confirmer par une étude plus fine des génomes des Bordetella.

8.4 Génomes des Gammaprotéobactéries

Les génomes multipartites des Gammaprotéobactéries sont trouvés chez tous les membres (décrits à ce jour) des Vibrionales et chez deux *Pseudoalteromonas*, parmi les Alteromonadales : *Pseudoalteromonas haloplanktis* TAC125 et *Pseudoalteromonas* sp. SM9913. Le génome d'*Aliivibrio fisheri* ES114 (Vibrionaceae, Vibrionales) est comparé aux génomes de *Vibrio cholerae* O1 biovar El Tor N16961 (Vibrionaceae, Vibrionales),

Photobacterium profundum SS9 (Photobacteriaceae, Vibrionales), *Shewanella amazonen*sis SB2B (Shewanellaceae), et *Pseudoalteromonas haloplanktis* TAC125 (Alteromonadaceae). En raison de la taille des génomes, Last est utilisé.



(A) A. fischeri ES114 (abscisse) vs. V. cholerae N16961 (ordonnée).



(B) A. fischeri ES114 (abscisse) vs. P. profundum SS9 (ordonnée).



(C) A. fischeri ES114 (abscisse) vs. P. haloplanktis TAC125 (ordonnée).



(D) A. fischeri ES114 (abscisse) vs. Shewanella amazonensis SB2B (ordonnée).



(E) V. cholerae N16961 (abscisse) vs. P. profundum SS9 (ordonnée).

FIGURE 8.11: Synténie entre Aliivibrio fischeri ES114 et Vibrio cholerae O1 biovar El Tor N16961 (8.11a), Photobacterium profundum SS9 (8.11b), Pseudoalteromonas haloplanktis TAC125 (8.11c) et Shewanella amazonensis SB2B (8.11d), et entre Vibrio cholerae O1 biovar El Tor N16961 et Photobacterium profundum SS9 (8.11e).

TABLE 8.10: Valeurs de l'indice synténique S_G (éq. 8.1) pour les réplicons du génome G d'Aliivibrio fischeri ES114 par rapport aux génomes G_{ref} de Vibrio cholerae O1 biovar El Tor N16961 (8.10a), Photobacterium profundum SS9 (8.10b), Pseudoalteromonas haloplanktis TAC125 (8.10c) et Shewanella amazonensis SB2B (8.10d, génome monopartite), et pour les réplicons de Vibrio cholerae O1 biovar El Tor N16961 par rapport au génome de Photobacterium profundum SS9 (8.10e).

(A) A. fischeri ES114 vs. V. cholerae N16961.

(B) A. fischeri ES114 vs. P. profundum SS9.

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	r_{pPBPR1}^{ref}
r_{chr}	100	100	r_{chr}	100	100	n.a.
r_{RECE}	10	544	r_{RECE}	11	226	n.a.

(C) A. fischeri ES114 vs. P. haloplanktis TAC125.

(D) A. fischeri ES114 vs. S. amazonensis SB2B.

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	$G \diagdown G_{ref}$	r_{chr}^{ref}
r_{chr}	100	100	r_{chr}	100
r_{RECE}	6	42	r_{RECE}	1

(E) V. cholerae N16961 vs. P. profundum SS9.

$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{RECE}^{ref}	r_{pPBPR1}^{ref}
r_{chr}	100	100	n.a.
r_{RECE}	6	311	n.a.

Curieusement, Il existe une synténie relativement plus forte entre les chromosomes et les RECE de *A. fisherii* et *P. profundum* par comparaison à *V. cholerae* (Figures 8.11d et 8.11e), bien qu'ils soient phylogéniquement moins proches qu'*A. fisherii* et *V. cholerae.* Il semble donc que des phénomènes de transfert entre les chromosomes et les RECE des *Aliivibrio* et *Photobacterium* ont eu lieu récemment.

8.5 Génomes des Cyanobactéries

Deux génomes multipartites sont actuellement référencés chez les Cyanobactéries et appartiennent à deux ordres différents.

8.5.1 Analyse des Nostocales

Le génome multipartite d'Anabaena sp. 90 est comparé au génome monopartite de Anabaena sp. PCC 7120 (Figure 8.13 et Table 8.11). blastn est utilisé.



FIGURE 8.12: Synténie d'Anabaena sp. 90 (ordonnée) vs. Anabaena sp. PCC 7120 (abscisse).

TABLE 8.11: Valeurs de l'indice synténique de synténie S_G (éq. 8.1) pour les réplicons du génome G d'Anabaena sp. 90 par rapport au génome G_{ref} d'Anabaena sp. PCC 7120.

	1120:								
$G \diagdown G_{ref}$	r_{chr}^{ref}	$r_{pCC7120\alpha}^{ref}$	$r_{pCC7120\alpha}^{ref}$	$r_{pCC7120\beta}^{ref}$	$r_{pCC7120\gamma}^{ref}$	$r_{pCC7120\delta}^{ref}$	$r_{pCC7120\epsilon}^{ref}$	$r_{pCC7120\zeta}^{ref}$	
r_{chr}	100	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	
r_{RECE}	120	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	

Comme pour le génome multipartite d'Asticcacaulis, les zones de synténie entre le RECE d'Anabaena sp. 90 et le chromosome d'Anabaena sp. PCC 7120 sont relativement plus grandes qu'entre les chromosomes de ces deux espèces. Par ailleurs, le RECE d'Anabaena sp. 90 ne possède aucune région de synténie avec les plasmides d'Anabaena sp. PCC 7120. Enfin, malgré un signal synténique plus faible, **des résultats similaires ont été obtenus en comparant Anabanea avec des génomes monopartites de cyanobactéries plus distantes phylogéniquement : Trichodesmium erythraeum IMS101 et Synechococcus elongatus PCC 7942** (résultats non montrés).

8.5.2 Analyse des Chroococcales

Le génome multipartite de *Cyanothece* sp. ATCC 51142 est comparé au génome monopartite de *Cyanothece* sp. PCC 8801 (Figure 8.13 et Table 8.12). *blastn* est utilisé.



FIGURE 8.13: Synténie entre *Cyanothece* sp. ATCC 51142 (ordonnée) et *Cyanothece* sp. PCC 8801 (abscisse).

TABLE 8.12: Valeurs de l'indice synténique de synténie S_G (éq. 8.1) pour les réplicons du génome G de Cyanothece sp. ATCC 51142 par rapport au génome G_{ref} de Cyanothece sp. PCC 8801.

$G \diagdown G_{ref}$	r_{chr}^{ref}	$r_{pP880101}^{ref}$	$r_{pP880102}^{ref}$	$r_{pP880103}^{ref}$
r_{chr}	100	n.a.	n.a.	n.a.
r_{RECE}	0	n.a.	n.a.	n.a.
$r_{plasmideA}$	0	n.a.	n.a.	n.a.
$r_{plasmideB}$	0	n.a.	n.a.	n.a.
$r_{plasmideC}^{ref}$	0	n.a.	n.a.	n.a.
$r_{plasmideD}^{ref}$	0	n.a.	n.a.	n.a.

Aucune synténie n'est detectée entre le RECE de *Cyanothece* sp. ATCC 51142 et le chromosome de *Cyanothece* sp. PCC 8801 ou ses plasmides. La synténie partagée entre les deux chromosomes est relativement faible pour des génomes de bactéries du même genre. Des résultats similaires ont été obtenus en utilisant d'autres génomes de *Cyanothece* spp. et *Acaryochloris marina* MBIC11017 (résultats non montrés). Les signaux synténiques étant très faibles, il est difficile de comparer les génomes de *Cyanothece* avec des génomes monopartites d'espèces plus distantes.

8.6 Génomes des Bacteroidetes

Deux génomes multipartites ont été identifiés parmi les *Prevotella*, espèces appartenant aux Bacteroidia, un des ordres de Bacteroidetes. Le génome multipartite de *P. intermedia* 17 a été comparé aux génomes monopartites de *P. ruminicola* 23, de *P. denticola* F0289 et de *Bacteroides fragilis* YCH46 (Figure 8.14 et Table 8.13). blastn a été utilisé. Étrangement, le chromosome (présence de *dnaA*) de *P. intermedia* 17 est le plus petit des réplicons essentiels du génome.

TABLE 8.13: Valeurs de l'indice synténique S_G (éq. 8.1) pour les réplicons du génome G de Prevotella intermedia 17 par rapport aux génomes G_{ref} de P. ruminicola 23 (8.13a), P. denticola F0289 (8.13b) et Bacteroides fragilis YCH46 (8.13c).

(A)	A) P. intermedia 17 vs. P. ru- minicola 23.		(B) P. intermedia 17 vs. P. den- ticola F0289.			((C) P. intermedia 17 vs. B. fra- gilis YCH46.		
	$G \diagdown G_{ref}$	r_{chr}^{ref}		$G \diagdown G_{ref}$	r_{chr}^{ref}		$G \diagdown G_{ref}$	r_{chr}^{ref}	r_{pBFY46}^{ref}
	r_{chr}	100		r_{chr}	100	_	r_{chr}	100	n.a.
	r_{RECE}	153		r_{RECE}	178		r_{RECE}	158	n.a.

Dans cette configuration, le RECE de *P. intermedia* 17 présente la plus forte synténie avec les chromosomes des génomes monopartites comparés. Au contraire, le réplicon annoté dans RefSeq comme chromosome chez *P. intermedia* 17 se rapproche davantage des RECE décrits précédemment que des chromosomes authentiques (*i.e.*, ressemblant plus aux chromosomes). La synténie de ce "chromosome-RECE" de *P. intermedia* 17 s'avère cependant relativement importante en comparaison aux chromosomes et, de plus, **est conservée pour des génomes plus éloignés appartenant aux Bacteroidetes**. Toutefois, une interprétation définitive n'est pas possible compte tenu du peu de zones de synténie conservées entre génomes du même genre (*Prevotella*).



(A) P. intermedia 17 (abscisse) vs. P. ruminicola 23 (ordonnée).



(B) P. intermedia 17 (abscisse) vs. P. denticola F0289 (ordonnée).



(C) P. intermedia 17 (abscisse) vs. B. fragilis YCH46 (ordonnée).

FIGURE 8.14: Synténie entre Prevotella intermedia 17 et P. ruminicola 23 (8.14a), P. denticola F0289 (8.14b), Bacteroides fragilis YCH46 (8.14c).

8.7 Discussion

Cette étude des synténies des génomes multipartites souligne les différences entre RECE et plasmides. Les RECE des génomes multipartites possèdent, avec les chromosomes des espèces phylogéniquement proches, des taux de synténie plus importants que les plasmides (cf. résultats pour les plasmides d'Asticcacaulis, Anabaena, Sphingobium, Dinoroseobacter et Burkholderia). Cette constatation est compatible avec soit une formation des RECE à partir de plasmides "enrichis" en gènes du chromosome par transfert(s) intra-génomique(s), soit une stabilisation sous forme de réplicon d'un fragment de chromosome (*cf.* Chapitre 2, §2.5). Cette observation a aussi été retrouvée dans le cas de certains plasmides que nous avons identifiés comme de potentiels RECE (notamment ceux de *Ruegeria* sp. TM1040 et d'*Azospirillum brasilense* Sp245) parmi les plasmides (*cf.* Chapitre 7), ce qui renforce l'hypothèse d'essentialité de ces réplicons.

Cependant, la structure génomique des RECE diverge plus rapidement que celle des chromosomes [Bavishi et al., 2010]. Les taux de synténie entre RECE et chromosomes ont donc été comparés pour des espèces ayant divergé à différents âges (sur la base de leurs relations phylogéniques). Pour de nombreux génomes multipartites, le degré de synténie relatif des RECE par rapport à celui des chromosomes s'atténue lorsque les génomes multipartites sont comparés aux génomes de plus en plus éloignés phylogéniquement. Les génomes multipartites des Vibrionales et des Burkholderiales en sont des exemples typiques. La structure génomique de ces espèces étant bien documentée (*cf.* Chapitre 2, Table 2.2), nos résultats confirment l'origine plasmidique des RECE présents dans ces groupes de bactéries.

L'étude des RECE des Rhizobiales, lignée également très étudiée, permet de rapprocher les synténies observées entre réplicons à différents événements génétiques : intégration du plasmide ancestral aux RECE, divergences évolutives entre les différents RECE et perte du plasmide ITR originel (ce chapitre, §8.2.3). De façon similaire, les résultats obtenus pour les RECE de *Rhodobacter* et de *Sphingobium* suggèrent que les deux espèces ont été soumises aux mêmes phénomènes génomiques.

Cependant, dans d'autres cas, *Asticcacaulis, Anabaena, Paracoccus et Prevotella*, le signal synténique entre RECE et chromosomes de ces espèces est relativement maintenu et important pour des espèces distantes. Ces RECE semblent de plus posséder une synténie relative (mesurée par l'indice synténique) supérieure à celles des autres RECE étudiés. Pour les RECE d'*Asticcacaulis* et d'*Anabaena*, on constate que la taille relative des régions synténiques des RECE avec les chromosomes d'espèces plus ou moins proches est plus importante que dans le cas de comparaisons entre chromosomes. Ces RECE sont donc en contradiction avec les processus génomiques communément utilisés pour caractériser les génomes multipartites :

- Une synténie conservée entre RECE et chromosome contredit l'hypothèse selon laquelle les RECE sont soumis à une évolution plus rapide que les chromosomes, et souligne la conservation et l'importance fonctionnelle de certaines régions de ces RECE.
- Une synténie importante du RECE, voire plus importante relativement que pour le chromosome, avec les chromosomes d'espèces mono- ou multipartites est peu compatible avec l'hypothèse d'"enrichissement" d'un plasmide par transfert(s) intragénomique(s). Elle est par contre plus cohérente avec celle d'une coupure du chromosome de l'espèce monopartite ancestrale. La structure des régions synténiques des réplicons d'Asticcacaulis laisse, de plus, penser que son chromosome et son RECE sont complémentaires, témoignant ainsi d'une scission originelle.

Les taux de variation des régions synténiques entre réplicons sont spécifiques à chaque groupe bactérien. Même si les RECE des Burkholderiales et des Vibrionales découlent vraisemblablement des mêmes processus génomiques, le taux de synténies entre leurs RECE et les chromosomes d'espèces monopartites est plus faible pour les RECE des Vibrionales témoignant de mécanismes génomiques spécifiques à ce groupe bactérien (évolution plus rapide du RECE, transferts intra-génomiques plus rares...). Les scores de l'indice synténique (éq. 8.1) peuvent refléter des artefacts génomiques et doivent être interprétés avec prudence. Le score élevé obtenu pour la comparaison du RECE de *Sphingobium* avec le chromosome de *Caulobacter* (Table 8.5a) est fortement influencé par la faible synténie observée entre les deux chromosomes et par la petite taille du RECE comparativement à celle du chromosome, et ne témoigne pas forcément d'une conservation de la synténie chez le RECE.

Enfin, chez certains groupes d'espèces telles que les cyanobactéries, les signaux de synténie inter-génomes sont trop faibles pour être interprétés. Chez d'autres groupes bactériens, trop peu de génomes sont actuellement disponibles pour permettre une comparaison cohérente. La confirmation des tendances mises en lumière par ces premiers résultats requiert une étude globale approfondie de la synténie inter-génomes avec, pour chaque catégorie de réplicons, un nombre de génomes.

Chapitre 9

Discussion générale

9.1 Méthodologies et principaux résultats

Les analyses des chapitres 5, 6 et 7 ont permis de mettre en évidence et de caractériser des biais entre les réplicons, selon leur type (chromosome, plasmide ou RECE), ou entre génomes mono- et multipartites, par rapport à leurs STIG (Figure 9.1).



FIGURE 9.1: Synthèse des résultats.

Ces biais ont été établis sur la base d'unités structurales, les clusters de protéines homologues, et fonctionnelles, par le choix des protéines selon leurs liens avec les STIG. Bien que la discrimination entre RECE, chromosomes et plasmides soit nette, celle entre génomes mono- et multipartites est plus floue. Parmi les RECE, différents types de biais sont observés : un biais fort caractérisé par la présence de nombreux gènes de type chromosomique, notamment détectable pour les RECE de *Prevotella, Anabaena, Asticcacaulis* et *Paracoccus*, et un biais moins marqué en comparaison des distributions
des gènes des STIG des plasmides, observable en particulier pour les RECE de *Deino*coccus, Leptospira et Cyanothece. Enfin, les critères fonctionnels (*i.e.*, les annotations) semblent permettre une meilleure discrimination que les critères structuraux (*i.e.*, les clusters de protéines homologues), suggérant que les **RECE sont mieux définis par** les fonctionnalités de leurs STIG que par l'histoire évolutive de leurs gènes.

9.1.1 Implications de la discrimination des RECE par les STIG

La présence sur un RECE d'un gène supplémentaire codant une fonction donnée a plusieurs implications possibles quant au rôle et au devenir de ce gène. Par exemple, un gène sur un RECE peut être soumis à de contraintes évolutivess moins fortes que sur le chromosome et ainsi, acquérir une nouvelle fonction plus facilement au sein de l'organisme (cf. Chapitre 2 §2.6). Un biais en gènes chromosomiques sur un RECE peut alors d'une part, être la conséquence d'échanges génomiques favorisés par la cohabitation durable entre RECE et chromosome et, d'autre part, refléter un mécanisme évolutif spécifique des génomes multipartites bactériens. Un biais en gènes plasmidiques sur les RECE, par rapport aux chromosomes, peut refléter les origines plasmidiques des RECE. Cependant, l'intégration des RECE dans le génome stable requiert le développement de mécanismes génétiques supplémentaires, dont certains ont déjà été caractérisés (cf. Chapitre 2 $\S2.4$). Ainsi, on peut légitimement se demander si les biais observés en gènes des STIG présents sur les RECE sont des conséquences des origines évolutives de ces éléments génomiques et de leur cohabitation avec le génome stable, ou bien reflètent l'adaptation des mécanismes des STIG des génomes multipartites permettant l'intégration des RECE dans le cycle cellulaire. Cidessous sont donnés divers arguments en faveur de cette seconde hypothèse :

- Des biais en gènes essentiels sur un ensemble de RECE ont déjà été caractérisés mais, contrairement aux gènes des STIG, aucune tendance globale n'a été identifiée [Harrison et al., 2010; Harrison, 2011].
- ► Il existe des mécanismes d'adaptation fonctionnelle communs à des génomes multipartites provenant de différentes lignées bactériennes, tels que l'adaptation du système parABS chez les Burkholdériales et chez V. cholerae (cf. Chapitre 2 §2.4) mais aussi chez certains plasmides des Rhizobiales, qui possédent un système repABC singulier. Ces systèmes sont présents de facon très majoritaire sur des mégaplasmides à bas taux de copies ayant un rôle important dans la fitness des cellules ainsi que sur certains RECE [Cervantes-Rivera et al., 2011; Petersen et al., 2013]. Il est alors raisonnable de supposer que ces systèmes contribuent à la stabilisation des réplicons accessoires dans le génome. Nos analyses permettent d'identifier un biais significatif en gènes codant des fonctions annotées parA/parB (cf. Chapitre 6 §6.4.3). Cette caractéristique est par exemple retrouvée sur les RECE des Leptospira ou de Deinococcus (cf. Chapitre 5 §5.3.2.1), laissant supposer que l'adaptation des systèmes parABS, ou de manière plus générale des systèmes de partition, est une condition sine qua non de la stabilisation des réplicons.
- ► Certains gènes ayant des annotations telles que celles correspondant aux initiateurs de la réplication chromosomique *dnaA*, ou encore à *ftsZ*, sont présents de façon quasi ubiquiste sur les chromosomes. Au contraire, les gènes codant pour les réplicateurs plasmidiques Rep se trouvent de facon très minoritaire sur les chromosomes. Ces

biais suggèrent alors que la présence de gènes des STIG sur un réplicon bactérien sont liés à des mécanismes des STIG spécifiques.

► Les chromosomes peuvent être classés selon la taxonomie de leur hôte en utilisant des marqueurs génétiques de type ARNr reflétant l'évolution verticale de ces réplicons. Un groupe taxonomique de chromosomes (par exemple une espèce) reflète alors les caractéristiques fonctionnelles et structurales communes de ces génomes. Contrairement aux chromosomes, les plasmides représentent des unités génomiques fonctionnelles et structurales qui sont soumises fortement aux THG et pour lesquelles il est plus difficile de trouver des critères génétiques de classification. Il a été suggéré qu'une partie des gènes des STIG, les recombinases, les systèmes de partition des réplicons et d'initiation de la réplication (*cf.* Chapitre 1 et Chapitre 2 §2.7.2), présents sur les plasmides pouvaient servir à classer ces réplicons dans différentes classes taxonomiques et fonctionnelles. Ces propositions se trouvent confirmées par les résultats des analyses de clustering où différents groupes de plasmides, liés à la taxonomie, sont retrouvés (*cf.* §2.7.2).

9.1.2 Études complémentaires

- ▶ Les biais des STIG mis en évidence et caractérisés dans cette étude ont révélé certains phénomènes qu'il est nécessaire de caractériser *via* des analyses complémentaires.
 - L'inclusion de davantage de fonctions liées directement ou indirectement aux STIG permettrait d'une part de mieux caractériser les génomes de certaines espèces pour lesquelles les modèles des STIG sont encore incompris et incomplets. D'autre part, certaines fonctions indirectes des STIG, telles que les réplicases ou les systèmes de conjugaison, n'ont pas été prises en considération dans notre étude. Les inclure dans un second temps peut apporter des éléments supplémentaires de comparaison. Enfin, l'inclusion des gènes liés au génome-cœur peut aussi permettre de comparer les biais pour ces gènes avec ceux observés pour les seuls STIG.
 - De même que pour à l'analyse des réplicons par la recherche d'homologues des STIG, il semble pertinent d'inclure dans l'analyse les motifs structurels dont le rôle est lié aux STIG (cf. Chapitre 1 §1.3.2). Cependant, l'identification et l'annotation de motifs soulèvent des difficultés techniques, et une expertise spécifique dans la détection de motifs est nécessaire du fait de leur taille réduite (≈ 50pb généralement) et de leur variabilité. Dans l'objectif de caractériser les sites dif des RECE des Prevotella, des analyses ont été menées parallèlement sans succès. En utilisant les sites dif déjà caractérisés de génomes d'espèces proches, il nous a été impossible d'identifier avec certitude les positions exactes des sites dif sur ces réplicons (résultats non montrés).
 - La structure de la séquence *ori* d'origine de réplication des réplicons bactériens organisée en interne en motifs fonctionnels et, sur le réplicon, à proximité de certains gènes refléte le type du réplicon (*cf.* Chapitre 1 §1.4.1.1). Inclure ces données dans l'analyse permettrait d'obtenir des paramètres supplémentaires liés à la stabilisation des réplicons dans le génome.

L'ensemble de ces études complémentaires peuvent alors servir à construire un modèle mixte, fondé sur des caractéristiques fonctionnelles et structurales, qui permettrait de mieux caractériser les réplicons bactériens.

- ▶ Par ailleurs, l'analyse de synténie du Chapitre 8 révèle certaines spécificités de la structure des RECE. Une analyse globale de synténie sur l'ensemble des réplicons bactériens devrait permettre de les affiner et de les caractériser précisément.
- ► Enfin, les analyses des Chapitres 3 et 7 montrent qu'il existe un biais des STIG présents entre génomes monopartites et génomes multipartites. Ces différences impliquent notamment des gènes liés à la régulation (de type *iciA* et *lrp*). Des études complémentaires, expérimentales ou analytiques, sont donc nécessaires pour confirmer ou infirmer l'importance de ces fonctionnalités et avancer dans la connaissance des processus cellulaires.

9.2 Remise en cause des hypothèses sur la nature des chromosomes secondaires

Bien que le concept de *chromid* soit intéressant pour décrire une partie des RECE (chez les Protéobactéries), différentes exceptions et contre-exemples existent.

Le critère d'identification du biais nucléotidique n'est pas universellement vrai. Chez *Cyanothece*, le RECE linéaire a un pourcentage en G+C de 38.6% comparé aux 37.9% du chromosome et 38.6%, 41.5%, 38.1% et 37.0% des plasmides additionnels [Welsh et al., 2008]. De même, le RECE de *Thermobaculum terrenum* a un pourcentage en G+C de 64% par opposition aux 48% du chromosome [Kiss et al., 2010]. Dans un dernier temps, le modèle du *chromid* n'autorise pas d'autre hypothèse évolutive que l'hypothèse **H2**, qui décrit l'apparition des RECE *via* l'enrichissement de plasmides originels (*cf.* Chapitre 2 §2.5).

Bien que des cas de formation de RECE relevant de l'hypothèse H1, expliquant la formation de RECE par la scission d'un chromosome originel, n'aient pas été formellement montrés, des données expérimentales de réarrangement génomique indiquent que cette modalité ne peut être rejetée. Le génome de *B. cereus* peut exister en un unique et large chromosome ou sous la forme d'un chromosome plus petit et de réplicons additionnels correspondant à des fragments du chromosome principal [Carlson and Kolsto, 1994]. Inversement, il a été montré que des réplicons accessoires peuvent intégrer le chromosome chez *S. meliloti* [Guo et al., 2003]. Les résultats concernant l'analyse par les STIG ainsi que les analyses de synténie des génomes d'*Asticcacaulis*, de *Deinococcus*, de *Prevotella* et d'*Anabaena* contrastent fortement avec les résultats obtenus avec les génomes multipartites "classiques", les plus étudiés et mieux caractérisés tels que ceux des Vibrionales et des Burkholdériales, et laissent envisager qu'ils sont issus d'un mécanisme de formation, autre que H2, compatible avec H1.

Enfin, étant donné que les génomes présentant des RECE sont disséminés dans l'ensemble du domaine bactérien, il est probable que les apparitions des génomes multipartites dans les différentes lignées bactériennes sont le fruit d'événements évolutifs distincts. Il est alors raisonnable de concevoir qu'au moins un organisme bactérien a acquis un second chromosome selon une autre modalité que celle proposée par l'hypothèse **H2**. La formation d'un RECE peut ne pas impliquer les plasmides comme vecteurs de réplicons additionnels. Le modèle du *chromid* développé par Harisson *et al.* n'est pas applicable dans l'ensemble des cas. La terminologie de "chromosome secondaire" est ainsi plus adaptée car elle ne préjuge pas d'un mode de formation particulier des RECE. Cependant, elle laisse présupposer qu'il existe un chromosome principal formé antérieurement aux RECE. Dans le cas de la formation de RECE selon les modalités de l'hypothèse **H1**, cette appellation n'est pas applicable car les deux chromosomes ont été formés de facon simultanée, lors de la scission du chromosome originel. La terminologie de **néo-chromosomes** apparaît alors mieux adaptée pour décrire les RECE et chromosome ainsi formés car elle place sur un pied d'égalité l'ensemble des réplicons essentiels néo-formés du génome.

9.3 Origine des biais de distribution des gènes des STIG

Plusieurs mécanismes génétiques peuvent être envisagés pour expliquer la singularité des distributions des gènes des STIG sur les RECE. Le résultats présentés dans le Chapitre 8 permettent de proposer des éléments de réponse.

- Transfert latéral intra-génomique. Ces mécanismes impliquent que les RECE ont acquis des gènes des chromosomes ou des plasmides *via* un échange intragénomique et que le gène transféré du réplicon "donneur" est présent *uniquement* sur le RECE. Ces phénomènes ont été décrits pour les génomes multipartites des *Burkholderia* et des Rhizobiales notamment (*cf.* Chapitre 8 §8.2.3 et §8.3).
- Duplication suivie de transfert intra-génomique. Ce mécanisme est similaire au précédent mais une copie du gène des STIG présent sur le RECE est aussi présent sur le réplicon "donneur", de façon analogue au mécanisme de paralogie. La présence d'une copie supplémentaire de gène annoté hfq sur les RECE des Burkholderia a été rapportée (cf. Chapitre 6 §6.4.3).
- Coupure du réplicon. Ce mécanisme implique la coupure d'un réplicon originel donnant naissance à deux réplicons fonctionnels. L'acquisition de STIG fonctionnels pour chacun des deux nouveaux réplicons passe par l'intégration de modules génétiques externes, par exemple plasmidiques. Il n'existe pas actuellement de recensement et de description précise de ce type de formation de réplicon chez les bactéries. Néanmoins, un tel mécanisme a été suggéré par divers auteurs (*cf.* Chapitre 2 §2.5), et *in vitro* chez *B. cereus*, le chromosome a été expérimentalement scindé en deux réplicons stables [Itaya and Tanaka, 1997]. Le fort biais en gènes des STIG observés sur certains RECE ainsi que les résultats des analyses de synténie (Chapitre 8) font supposer que les RECE de Prevotella, Paracoccus, Asticcacaulis et Anabaena ont été formés selon cette modalité.
- Transfert latéral inter-génomique. L'acquisition de gènes des STIG étrangers au génome via des THG est un mécanisme à l'origine des biais observés sur les RECE. Une partie des gènes présents sur un réplicon extra-chromosomique d'Azospirillum brasilense CBG497 caractérisés comme potentiels RECE, sont des orthologues de gènes de génomes d'espèces, genres ou phylum différents (xénologues), dont deux codent une transposase [Acosta-Cruz et al., 2012].

• Fusion de deux réplicons. Dans certains cas, un réplicon peut s'intégrer au sein d'un autre. Ces observations ont été faites *in vitro*, chez *E. coli* [Bernander et al., 1991] par exemple et *in vivo* où il a notamment été reporté que certains RECE pouvaient intégrer de facon stable le chromosome [Guo et al., 2003]. Récemment, il a été montré *in vitro* que la réplication du RECE de *Vibrio cholerae* sans l'intervention de Dam et RtcB est possible par la fusion du RECE avec le chromosome, la réplication s'effectuant *via* la machinerie moléculaire du chromosome [Val et al., 2014].

9.4 Proposition d'un modèle moléculaire d'origine des néochromosomes

On peut alors supposer que ces différents mécanismes génétiques sont grandement impliqués dans la diversification des STIG chez les réplicons bactériens. Ils peuvent par exemple contribuer à la *replicon takeover hypothesis*, selon laquelle un réplicon accessoire peut rendre fonctionnelle sa machinerie réplicative par intégration au sein d'un autre réplicon, chromosome ou plasmide [McGeoch and Bell, 2008]. Par exemple, l'intégration *in vitro* d'un plasmide R1 au sein du chromosome d'*E. coli* a pu rendre inutile l'action de DnaA, la réplication se faisant *via* l'origine et les initiateurs du plasmide intégré [Bernander et al., 1991]. On peut aussi rapprocher de cette hypothèse le cas des réplicons contenant de multiples origines de réplication actives (ou plus exactement, des éléments génomiques à multiple réplicons), comme par exemple chez certains plasmides (*cf.* §1.2) qui bénéficient ainsi d'un large spectre d'hôtes [Toukdarian, 2004].

Des mécanismes de régulation sont à l'oeuvre entre le chromosome et le mégaplasmide chez Streptomyces clavuligerus (Actinobactéries) et font vraisemblablement intervenir des protéines plasmidiques héritées du chromosome via des mécanismes de recombinaison ou de transposition [Medema et al., 2010]. Le taux élevé de gènes de type recA, codant pour une protéine impliquée dans la réparation de l'ADN, dans le génome d'Acaryochloris marina (Cyanobactéries) a probablement contribué à l'expansion de son génome, ces gènes étant apparus par duplication et/ou THG [Swingley et al., 2008]. Enfin, un dernier exemple intéressant est l'hypothèse selon laquelle certains systèmes de conjugaison seraient d'abord apparus chez les Protéobactéries puis auraient été diffusés horizontalement dans l'ensemble des lignées bactériennes ainsi que chez les Archées [Guglielmini et al., 2013]. En particulier, TraB, protéine appartenant aux systèmes de sécrétion de type IV, et FtsK, protéine majeure du fonctionnement cellulaire liant la réplication des réplicons à la division cellulaire, dériveraient d'une protéine ancestrale commune [Vogelmann and Ammelburg, 2011].

La distribution des gènes des STIG sur les réplicons bactériens forment ainsi un marqueur pertinent de la stabilisation et de l'intégration des réplicons dans le génome bactérien et témoigne des spécificités des réplicons. Le passage d'une architecture génomique monopartite à une architecture multipartite requiert des STIG spécifiques, que notre étude des caractéristiques discriminantes des RECE a mis en évidence. En considérant l'ensemble des génomes, il semble de plus que le bon fonctionnement des génomes multipartites par rapport aux génomes monopartites réclame le soutien de gènes supplémentaires, impliqués dans la régulation (régulateurs de type iciA ou lrp) ou intervenant dans la partition.

9.5 Continuité du matériel génomique

L'exemple de la propagation de systèmes de conjugaison à travers l'ensemble des lignées bactériennes à partir d'un système originel limité initialement à un seul groupe d'espèce illustre bien la notion de continuité du matériel génomique. Certains indices génomiques montrent que des protéines similaires liées aux STIG, vraisemblablement héritées de protéines ancestrales communes, existent chez les Archées, Eucaryotes et Bactéries [Mc-Geoch and Bell, 2008]. Parallèlement à l'évolution des systèmes génétiques, l'étude des réplicons bactériens permet de mettre en avant la continuité du matériel génomique. Chez les réplicons de *B. subtilis* ainsi que dans d'autres génomes bactériens, des mégaplasmides ont été formés à partir de la fusion de petits plasmides, ce mécanisme pouvant être une cause majeure de formation de mégaplasmide chez les bactéries [Zheng et al., 2013]. Il existe alors non seulement une plasticité génétique des réplicons mais aussi de leur degré de stabilisation dans le génome. Tout comme plasmides et chromosomes, les RECE et néo-chromosomes sont des espèces génomiques marqueurs de processus spécifiques d'intégration et de complexification (ou *dés*intégration et réduction) des génomes bactériens.

Bibliographie

- Acosta-Cruz, E., F. Wisniewski-Dyé, Z. Rouy, V. Barbe, M. Valdés, and P. Mavingui (2012). Insights into the 1.59-Mbp largest plasmid of *Azospirillum brasilense* CBG497. *Archives of microbiology* 194(9), 725–736.
- Agrawal, R. and R. Srikant (1994). Fast algorithms for mining association rules. In Proceedings of the 20th Very Large Data Bases (VLDB) International Conference, Santiago, Chile, Volume 1215, pp. 487–499.
- Allardet-Servent, A., S. Michaux-Charachon, E. Jumas-Bilak, L. Karayan, and M. Ramuz (1993). Presence of one linear and one circular chromosome in the Agrobacterium tumefaciens C58 genome. Journal of Bacteriology 175(24), 7869–7874.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215(3), 403–410.
- Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman (1997). Gapped BLAST and PSI-BLAST : a new generation of protein database search programs. *Nucleic Acids Research* 25(17), 3389–3402.
- Amadou, C., G. Pascal, S. Mangenot, M. Glew, C. Bontemps, D. Capela, S. Carrère, S. Cruveiller, C. Dossat, A. Lajus, et al. (2008). Genome sequence of the βrhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Research* 18(9), 1472–1483.
- Andreopoulos, B., A. An, X. Wang, and M. Schroeder (2009). A roadmap of clustering algorithms : finding a match for a biomedical application. *Briefings in Bioinforma*tics 10(3), 297–314.
- Apeltsin, L., J. H. Morris, P. C. Babbitt, and T. E. Ferrin (2011). Improving the quality of protein similarity network clustering algorithms using the network edge weight distribution. *Bioinformatics* 27(3), 326–333.
- Arlot, S. and A. Celisse (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys* 4, 40–79.
- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, and G. Sherlock (2000). Gene ontology : tool for the unification of biology. *Nature Genetics* 25(1), 25–29.

- Azam, T. A. and A. Ishihama (1999). Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *Journal of Biological Chemistry* 274 (46), 33105–33113.
- Baek, J. H. and D. K. Chattoraj (2014). Chromosome I controls chromosome II replication in Vibrio cholerae. PLoS Genetics 10(2), e1004184.
- Bapteste, E., M. A. O'Malley, R. G. Beiko, M. Ereshefsky, J. P. Gogarten, L. Franklin-Hall, F.-J. Lapointe, J. Dupré, T. Dagan, Y. Boucher, and W. Martin (2009). Prokaryotic evolution and the tree of life are two different things. *Biology Direct* 4(1), 34.
- Barbe, V., M. Bouzon, S. Mangenot, B. Badet, J. Poulain, B. Segurens, D. Vallenet, P. Marlière, and J. Weissenbach (2011). Complete genome sequence of *Streptomyces cattleya* NRRL 8057, a producer of antibiotics and fluorometabolites. *Journal of Bacteriology* 193(18), 5055–5056.
- Baril, C., J. Herrmann, C. Richaud, D. Margarita, and I. Girons (1992). Scattering of the rRNA genes on the physical map of the circular chromosome of *Leptospira interrogans* serovar *icterohaemorrhagiae*. Journal of Bacteriology 174(23), 7566–7571.
- Baril, C., C. Richaud, G. Baranton, and I. S. Girons (1989). Linear chromosome of Borrelia burgdorferi. Research in Microbiology 140(7), 507–516.
- Barnes, M. H., W. A. LaMarr, and K. A. Foster (2003). DNA gyrase and DNA topoisomerase of *Bacillus subtilis* : expression and characterization of recombinant enzymes encoded by the gyrA, gyrB and parC, parE genes. Protein Expression and Purification 29(2), 259–264.
- Barnett, M. J., R. F. Fisher, T. Jones, C. Komp, A. P. Abola, F. Barloy-Hubler, L. Bowser, D. Capela, F. Galibert, J. Gouzy, et al. (2001). Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proceedings of* the National Academy of Sciences U.S.A. 98(17), 9883–9888.
- Barre, F.-X., M. Aroyo, S. D. Colloms, A. Helfrich, F. Cornet, and D. J. Sherratt (2000). FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation. *Genes & Development* 14(23), 2976–2988.
- Barre, F. X. and D. Sherratt (2005). Chromosome dimer resolution. In N. P. Higgins (Ed.), *The Bacterial Chromosome*, ASM Press, Washington, DC. Pp. 513–524.
- Bastian, M., S. Heymann, and M. Jacomy (2009). Gephi : An open source software for exploring and manipulating networks. In *Third International AAAI Conference on* Weblogs and Social Media, San Jose Mc Enery Convention Center, May 17, 2009 – May 20, 2009. AAAI Publications.
- Bavishi, A., A. Abhishek, L. Lin, and M. Choudhary (2010). Complex prokaryotic genome structure : rapid evolution of chromosome II. *Genome* 53(9), 675–687.
- Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler (2008). GenBank. Nucleic Acids Research 36(Database issue), D25–D30.

- Bentley, S. D. and J. Parkhill (2004). Comparative genomic structure of prokaryotes. Annual Review of Genetics 38, 771–791.
- Bernander, R., S. Dasgupta, and K. Nordström (1991). The *E. coli* cell cycle and the plasmid R1 replication cycle in the absence of the DnaA protein. *Cell* 64(6), 1145–1153.
- Bigot, S., O. A. Saleh, C. Lesterlin, M. El Karoui, C. Dennis, M. Grigoriev, J.-F. Allemand, F.-X. Barre, F. Cornet (2005). KOPS : DNA motifs that control *E. coli* chromosome segregation by orienting the FtsK translocase. *The EMBO Journal* 24 (21), 3770–3780.
- Bigot, S., V. Sivanathan, C. Possoz, F.-X. Barre, and F. Cornet (2007). FtsK, a literate chromosome segregation machine. *Molecular Microbiology* 64 (6), 1434–1441.
- Blanca-Ordóñez, H., J. J. Oliva-García, D. Pérez-Mendoza, M. J. Soto, J. Olivares, J. Sanjuán, and J. Nogales (2010). pSymA-dependent mobilization of the *Sinorhizo*bium meliloti pSymB megaplasmid. Journal of Bacteriology 192(23), 6309–6312.
- Boekhorst, J. and B. Snel (2007). Identification of homologs in insignificant blast hits by exploiting extrinsic gene properties. *BMC Bioinformatics* 8, 356.
- Boyd, E. F., S. Almagro-Moreno, and M. A. Parent (2009). Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends in Microbiology* 17(2), 47– 53.
- Brantl, S. (2004). Plasmid replication control by antisense RNAs. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, ASM Press, Washington, DC. Pp. 47–62.
- Breiman, L. (2001). Random forests. Machine Learning 45(1), 5–32.
- Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984). *Classification and regression trees.* CRC press.
- Brilli, M., M. Fondi, R. Fani, A. Mengoni, L. Ferri, M. Bazzicalupo, and E. G. Biondi (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria : a comparative genomic analysis. *BMC Systems Biology* 4(1), 52.
- Brohée, S. and J. van Helden (2006). Evaluation of clustering algorithms for proteinprotein interaction networks. *BMC Bioinformatics* 7, 488.
- Browning, D. F., D. C. Grainger, and S. J. Busby (2010). Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Current Opinion in Microbiology* 13(6), 773–780.
- Bulach, D. M., R. L. Zuerner, P. Wilson, T. Seemann, A. McGrath, P. A. Cullen, J. Davis, M. Johnson, E. Kuczek, D. P. Alt, B. D. Peterson-Burch, R. L. Coppel, J. I. Rood, J. K. Davies, B. Adler (2006). Genome reduction in Leptospira borgpetersenii reflects limited transmission potential. Proceedings of the National Academy of Sciences U.S.A. 103(39), 14560–14565.

- Cairns, J. (1963). The chromosome of Escherichia coli. Cold Spring Harbor Symposia on Quantitative Biology 28, 43–46.
- Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden (2009). BLAST+ : architecture and applications. *BMC Bioinformatics* 10(1), 421.
- Carlson, C. R. and A.-B. Kolsto (1994). A small (2.4 Mb) Bacillus cereus chromosome corresponds to a conserved region of a larger (5.3 Mb) Bacillus cereus chromosome. Molecular Microbiology 13(1), 161–169.
- Carnoy, C. and C.-A. Roten (2009). The *dif*/Xer recombination systems in proteobacteria. *PloS one* 4(9), e6531.
- Casjens, S. (1998). The diverse and dynamic structure of bacterial genomes. Annual Review of Genetics 32(1), 339–377.
- Castillo-Ramírez, S., J. F. Vázquez-Castellanos, V. González, and M. A. Cevallos (2009). Horizontal gene transfer and diverse functional constrains within a common replication-partitioning system in Alphaproteobacteria : the repABC operon. BMC Genomics 10, 536.
- Cervantes-Rivera, R., F. Pedraza-López, G. Pérez-Segura, and M. A. Cevallos (2011). The replication origin of a repABC plasmid. *BMC Microbiology* 11(1), 158.
- Chaconas, G. and C. Chen (2005). Replication of linear bacterial chromosomes : No longer going around the circle. In N. P. Higgins (Ed.), *The Bacterial Chromosome*, ASM Press, Washington, DC. Pp. 525–539.
- Chain, P. S., D. M. Lang, D. J. Comerci, S. A. Malfatti, L. M. Vergez, M. Shin, R. A. Ugalde, E. Garcia, and M. E. Tolmasky (2011). Genome of *Ochrobactrum anthropi* ATCC 49188T, a versatile opportunistic pathogen and symbiont of several eukaryotic hosts. *Journal of Bacteriology* 193(16), 4274–4275.
- Chen, C.-Y., K.-M. Wu, Y.-C. Chang, C.-H. Chang, H.-C. Tsai, T.-L. Liao, Y.-M. Liu, H.-J. Chen, A. B.-T. Shen, J.-C. Li, et al. (2003). Comparative genome analysis of Vibrio vulnificus, a marine pathogen. Genome Research 13(12), 2577–2587.
- Chikina, M. D. and O. G. Troyanskaya (2011). Accurate quantification of functional analogy among close homologs. *PLoS Computational Biology* 7(2), e1001074.
- Choudhary, M., Y.-X. Fu, C. Mackenzie, and S. Kaplan (2004). DNA sequence duplication in *Rhodobacter sphaeroides* 2.4.1 : Evidence of an ancient partnership between chromosomes I and II. *Journal of Bacteriology* 186(7), 2019–2027.
- Choudhary, M., X. Zanhua, Y. X. Fu, and S. Kaplan (2007). Genome analyses of three strains of *Rhodobacter sphaeroides* : evidence of rapid evolution of chromosome II. *Journal of Bacteriology* 189(5), 1914–1921.

- Cicmil, N. (2008). Crystallization and preliminary X-ray crystallographic characterization of TrmFO, a folate-dependent tRNA methyltransferase from *Thermotoga maritima*. Acta Crystallographica Section F : Structural Biology and Crystallization Communications 64(3), 193–195.
- Clark, D. J. and O. Maaløe (1967). DNA replication and the division cycle in *Escherichia coli*. Journal of Molecular Biology 23(1), 99–112.
- Cooper, V. S., S. H. Vohr, S. C. Wrocklage, and P. J. Hatcher (2010). Why genes evolve faster on secondary chromosomes in Bacteria. *PLoS Computational Biology* 6(4), e1000732.
- Copley, S. D., J. Rokicki, P. Turner, H. Daligault, M. Nolan, and M. Land (2012). The whole genome sequence of *Sphingobium chlorophenolicum* L-1 : insights into the evolution of the pentachlorophenol degradation pathway. *Genome Biology and Evolution* 4(2), 184–198.
- Corcoran, C. P. and C. J. Dorman (2009). DNA relaxation-dependent phase biasing of the *fim* genetic switch in *Escherichia coli* depends on the interplay of H-NS, IHF and LRP. *Molecular Microbiology* 74(5), 1071–1082.
- Cornet, F. and M. Chandler (2004). Non-homologous recombination. In R. V. Miller and M. J. Day (Eds.), *Microbial Evolution : Gene Establishment, Survival, and Exchange*, ASM Press Washington, DC. Pp. 36–66.
- Cortes, C. and V. Vapnik (1995). Support-vector networks. *Machine Learning* 20(3), 273–297.
- Cortez, D., S. Quevillon-Cheruel, S. Gribaldo, N. Desnoues, G. Sezonov, P. Forterre, and M.-C. Serre (2010). Evidence for a Xer/dif system for chromosome resolution in archaea. *PLoS genetics* 6(10), e1001166.
- Coscia, M., F. Giannotti, and D. Pedreschi (2011). A classification for community discovery methods in complex networks. *Statistical Analysis and Data Mining* 4(5), 512–546.
- Croll, D. and B. A. Mcdonald (2012). The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathogens* 8(4), 8–10.
- Csardi, G. and T. Nepusz (2006). The igraph software package for complex network research. *InterJournal, Complex Systems* (1695).
- Darling, A. C., B. Mau, F. R. Blattner, and N. T. Perna (2004). Mauve : multiple alignment of conserved genomic sequence with rearrangements. *Genome Research* 14(7), 1394–1403.
- Das, B., E. Martínez, C. Midonet, and F.-X. Barre (2013). Integrative mobile elements exploiting Xer recombination. *Trends in Microbiology* 21(1), 23–30.
- Datta, S. and S. Datta (2006). Methods for evaluating clustering algorithms for gene expression data using a reference set of functional classes. *BMC Bioinformatics* 7, 397.

- de Hoon, M. J., S. Imoto, J. Nolan, and S. Miyano (2004). Open source clustering software. *Bioinformatics* 20(9), 1453–1454.
- de la Cueva-Méndez, G. and B. Pimentel (2007). Gene and cell survival : lessons from prokaryotic plasmid R1. *EMBO Reports* 8(5), 458–464.
- Del Solar, G., J. C. Alonso, M. Espinosa, and R. Díaz-Orejas (1996). Broad-host-range plasmid replication : an open question. *Molecular Microbiology* 21(4), 661–666.
- Del Solar, G., R. Giraldo, M. J. Ruiz-Echevarría, M. Espinosa, and R. Díaz-Orejas (1998). Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews* 62(2), 434–464.
- DelVecchio, V. G., V. Kapatral, R. J. Redkar, G. Patra, C. Mujer, T. Los, N. Ivanova, I. Anderson, A. Bhattacharyya, A. Lykidis, G. Reznik, L. Jablonski, N. Larsen, M. D'Souza, A. Bernal, M. Mazur, E. Goltsman, E. Selkov, P. H. Elzer, S. Hagius, D. O'Callaghan, J. J. Letesson, R. Haselkorn, N. Kyrpides and R. Overbeek (2002). The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. *Proceedings of the National Academy of Sciences U.S.A. 99*(1), 443–448.
- Demarre, G. and D. K. Chattoraj (2010). DNA adenine methylation is required to replicate both *Vibrio cholerae* chromosomes once per cell cycle. *PLoS Genetics* 6(5), e1000939.
- Dillon, S. C. and C. J. Dorman (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Reviews Microbiology* 8(3), 185–195.
- Doolittle, W. F. and E. Bapteste (2007). Pattern pluralism and the Tree of Life hypothesis. *Proceedings of the National Academy of Sciences U.S.A.* 104(7), 2043–2049.
- Dubarry, N., F. Pasta, and D. Lane (2006). ParABS systems of the four replicons of Burkholderia cenocepacia : new chromosome centromeres confer partition specificity. Journal of Bacteriology 188(4), 1489–1496.
- Duby, C. and S. Robin (2006). Analyse en composantes principales. *Institut National* Agronomique, Paris-Grignon 80.
- Duigou, S., K. G. Knudsen, O. Skovgaard, E. S. Egan, A. Løbner-Olesen, and M. K. Waldor (2006). Independent control of replication initiation of the two Vibrio cholerae chromosomes by DnaA and RctB. Journal of Bacteriology 188(17), 6419–6424.
- Ebersbach, G. and K. Gerdes (2005). Plasmid segregation mechanisms. Annual Review of Genetics 39(1), 453–479.
- Eddy, S. R. (1998). Profile hidden Markov models. Bioinformatics 14(9), 755-763.
- Eddy, S. R. (2011). Accelerated Profile HMM Searches. PLoS Computational Biology 7(10), e1002195.
- Eddy, S. R., T. J. Wheeler, and J. Farm (2013). HMMER User's Guide.

- Edgar, R. C. (2004). Muscle : multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32(5), 1792–1797.
- Edwards, D. H. and J. Errington (1997). The *Bacillus subtilis* DivIVA protein targets to the division septum and controls the site specificity of cell division. *Molecular Microbiology* 24(5), 905–915.
- Egan, E. (2005). Divided genomes : negotiating the cell cycle in prokaryotes with multiple chromosomes. *Molecular Microbiology* 56(5), 1129–1138.
- Egan, E. S., S. Duigou, and M. K. Waldor (2006). Autorepression of RctB, an initiator of Vibrio cholerae chromosome II replication. Journal of Bacteriology 188(2), 789–793.
- Egan, E. S. and M. K. Waldor (2003). Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* 114(4), 521–530.
- Enright, A. J., S. Van Dongen, and C. A. Ouzounis (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* 30(7), 1575–1584.
- Ester, M., H.-p. Kriegel, J. Sander, and X. Xu (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In E. Simoudis, J. Han, and U. M. Fayyad (Eds.), Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), Volume 1996, AAAI Press. Pp. 226–231.
- Fang, G., N. Bhardwaj, R. Robilotto, and M. B. Gerstein (2010). Getting started in gene orthology and functional analysis. *PLoS Computational Biology* 6(3), 8.
- Fang, Y. and J. Wang (2012). Selection of the number of clusters via the bootstrap method. *Computational Statistics & Data Analysis* 56(3), 468–477.
- Fernández-López, R., M. P. Garcillán-Barcia, C. Revilla, M. Lázaro, L. Vielva, and F. De La Cruz (2006). Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. *FEMS Microbiology Reviews* 30(6), 942–966.
- Finn, R. D., A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. L. Sonnhammer, J. Tate and M. Punta (2014). Pfam : the protein families database. *Nucleic acids research* 42(Database issue), D222–D230.
- Finn, R. D., J. Clements, and S. R. Eddy (2011). HMMER web server : interactive sequence similarity searching. *Nucleic Acids Research 39* (Web Server issue), W29– W37.
- Fitch, W. M. (1970). Distinguishing homologous from analogous proteins. Systematic Biology 19(2), 99–113.
- Fitch, W. M. (2000). Homology : a personal view on some of the problems. Trends in Genetics 16(5), 227–231.
- Francia, M. V., A. Varsaki, M. P. Garcillán-Barcia, A. Latorre, C. Drainas, and F. de la Cruz (2004). A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiology Reviews 28*(1), 79–100.

- Frech, C. and N. Chen (2010). Genome-wide comparative gene family classification. PLoS ONE 5(10), 14.
- Funnell, B. E. E. and R. A. Slavcev (2004). Partition systems of bacterial plasmids. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, ASM Press, Washington, DC. Pp. 81–103.
- Galardini, M., F. Pini, M. Bazzicalupo, E. G. Biondi, and A. Mengoni (2013). Replicondependent bacterial genome evolution : the case of *Sinorhizobium meliloti*. *Genome Biology and Evolution* 5(3), 542–558.
- Gan, G., C. Ma, and J. Wu (2007). Data clustering : theory, algorithms, and applications, Volume 20 of ASA-SIAM Series on Statistics and Applied Probability. SIAM, Society for Industrial and Applied Mathematics.
- Garcia Costas, A. M., Z. Liu, L. P. Tomsho, S. C. Schuster, D. M. Ward, and D. A. Bryant (2012). Complete genome of *Candidatus* Chloracidobacterium thermophilum, a chlorophyll-based photoheterotroph belonging to the phylum Acidobacteria. *Environmental Microbiology* 14(1), 177–190.
- Garcillán-Barcia, M. P., M. V. Francia, and F. De La Cruz (2009). The diversity of conjugative relaxases and its application in plasmid classification. *FEMS microbiology reviews* 33(3), 657–687.
- Geurts, P., D. Ernst, and L. Wehenkel (2006). Extremely randomized trees. Machine Learning 63(1), 3–42.
- Ghosh, S. K., S. Hajra, A. Paek, and M. Jayaram (2006). Mechanisms for chromosome and plasmid segregation. *Annual Review of Biochemistry* 75(1), 211–241.
- Glass, J. I., N. Assad-Garcia, N. Alperovich, S. Yooseph, M. R. Lewis, M. Maruf, C. A. Hutchison, H. O. Smith, and J. C. Venter (2006). Essential genes of a minimal bacterium. Proceedings of the National Academy of Sciences of the U.S.A. 103(2), 425–430.
- Goodner, B., G. Hinkle, S. Gattung, N. Miller, M. Blanchard, B. Qurollo, B. S. Goldman,
 Y. Cao, M. Askenazi, C. Halling, L. Mullin, K. Houmiel, J. Gordon, M. Vaudin,
 O. Iartchouk, A. Epp, F. Liu, C. Wollam, M. Allinger, D. Goughty, C. Scott,
 C. Lappas, B. Markelz, C. Flanagan, C. Crowell, J. Gurson, C. Lomo, C. Sear,
 G. Strub, C. Cielo, and S. Slater (2001). Genome sequence of the plant pathogen and
 biotechnology agent Agrobacterium tumefaciens C58. Science 294 (5550), 2323–2328.
- Graham, J. E., V. Sivanathan, D. J. Sherratt, and L. K. Arciszewska (2010). FtsK translocation on DNA stops at XerCD-*dif*. *Nucleic Acids Research* 38(1), 72–81.
- Guglielmini, J., F. de la Cruz, and E. P. C. Rocha (2013). Evolution of conjugation and type IV secretion systems. *Molecular Biology and Evolution* 30(2), 315–331.
- Guo, F., L. Ning, J. Huang, H. Lin, and H. Zhang (2010). Chromosome translocation and its consequence in the genome of *Burkholderia cenocepacia* AU-1054. *Biochemical* and *Biophysical Research Communications* 403(3–4), 375–379.

- Guo, X., M. Flores, P. Mavingui, S. I. Fuentes, G. Hernández, G. Dávila, and R. Palacios (2003). Natural genomic design in *Sinorhizobium meliloti* : novel genomic architectures. *Genome Research* 13(8), 1810–1817.
- Haas, B. J., A. L. Delcher, J. R. Wortman, and S. L. Salzberg (2004). DAGchainer : a tool for mining segmental genome duplications and synteny. *Bioinformatics* 20(18), 3643–3646.
- Haft, D. H., J. D. Selengut, and O. White (2003). The TIGRFAMs database of protein families. *Nucleic acids research* 31(1), 371–373.
- Hahn, M. W., E. Lang, U. Brandt, Q. L. Wu, and T. Scheuerl (2009). Emended description of the genus *Polynucleobacter* and the species *Polynucleobacter necessarius* and proposal of two subspecies, *P. necessarius* subsp. necessarius subsp. nov. and *P. necessarius* subsp. asymbioticus subsp. nov. *International Journal of Systematic and Evolutionary Microbiology 59*(8), 2002–2009.
- Hallet, B., V. Vanhooff, and F. Cornet (2004). DNA site-specific resolution systems. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, ASM Press, Washington, DC. Pp. 157–180.
- Hamel, L. H. (2011). Knowledge Discovery with Support Vector Machines, Volume 3. John Wiley & Sons.
- Hamoen, L. W., J.-C. Meile, W. De Jong, P. Noirot, and J. Errington (2006). SepF, a novel FtsZ-interacting protein required for a late step in cell division. *Molecular Microbiology* 59(3), 989–999.
- Han, J., M. Kamber, and J. Pei (2012). Data Mining : Concepts and Techniques, Third Edition. The Morgan Kaufmann Series in Data Management Systems. Morgan kaufmann, Elsevier.
- Han, J.-I., H.-K. Choi, S.-W. Lee, P. M. Orwin, J. Kim, S. L. LaRoe, T.-g. Kim, J. O'Neil, J. R. Leadbetter, S. Y. Lee, J. R. Leadbetter, S. Y. Lee, C. -G. Hur, J. C. Spain, G. Ovchinnikova, L. Goodwin, and C. Han (2011). Complete genome sequence of the metabolically versatile plant growth-promoting endophyte Variovorax paradoxus S110. Journal of Bacteriology 193(5), 1183–1190.
- Han, K., Z.-f. Li, R. Peng, L.-p. Zhu, T. Zhou, L.-g. Wang, S.-g. Li, X.-b. Zhang, W. Hu, Z.-h. Wu, N. Qin, and Y. -z. Lia (2013). Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Scientific Reports* 3, 2101.
- Harrison, P., R. Lower, N. Kim, and J. Young (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends in Microbiology* 16(4), 141–148.
- Harrison, P. W. (2011). Bacterial chromids are neither chromosomes nor plasmids. Ph.D. thesis, University of York.
- Hartigan, J. A. (1975). Clustering algorithms. John Wiley & Sons, Inc.

- Hartigan, J. A. and M. A. Wong (1979). Algorithm AS 136 : A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics) 28(1), 100–108.
- He, M., M. Sebaihia, T. D. Lawley, R. A. Stabler, L. F. Dawson, M. J. Martin, K. E. Holt, H. M. Seth-Smith, M. A. Quail, R. Rance, K. Brooks, C. Churcher, D. Harris, S. D. Bentley, C. Burrows, L. Clark, C. Corton, V. Murray, G. Rose, S. Thurston, A. van Tonder, D. Walker, B. W. Wren, G. Dougan, and J. Parkhill (2010). Evolutionary dynamics of *Clostridium difficile* over short and long time scales. *Proceedings of the National Academy of Sciences U.S.A.* 107(16), 7527–7532.
- Heidelberg, J. F., J. A. Eisen, W. C. Nelson, R. A. Clayton, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, L. Umayam, S. R. Gill, K. E. Nelson, T. D. Read, H. Tettelin, D. Richardson, M. D. Ermolaeva, J. Vamathevan, S. Bass, H. Qin, I. Dragoi, P. Sellers, L. McDonald, T. Utterback, R. D. Fleishmann, W. C. Nierman, O. White, S. L. Salzberg, H. O. Smith, R. R. Colwell, J. J. Mekalanos, J. C. Venter, and C. M. Fraser (2000). DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae. Nature 406 (6795), 477–483.
- Helinski, D. R. (2004). Introduction to plasmids : a selective view of their history. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, ASM Press, Washington, DC. Pp. 1–21.
- Hennig, C. (2007). Cluster-wise assessment of cluster stability. Computational Statistics & Data Analysis 52(1), 258–271.
- Hennig, C. (2008). Dissolution point and isolation robustness : Robustness criteria for general cluster analysis methods. *Journal of Multivariate Analysis 99*(6), 1154–1176.
- Henriques, A. O., P. Glaser, P. J. Piggot, and C. P. Moran Jr (1998). Control of cell shape and elongation by the rodA gene in *Bacillus subtilis*. *Molecular Microbiology* 28(2), 235–247.
- Heuer, H., Z. Abdo, and K. Smalla (2008). Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. *FEMS microbiology ecology* 65(3), 361–371.
- Higgins, N. P. (2005). The bacterial chromosome. ASM Press Washington, DC.
- Higgins, N. P. and A. Vologodskii (2004). Topological behavior of plasmid DNA. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, pp. 181–201. ASM Press, Washington, DC. Pp. 181–201.
- Hinnebusch, J. and K. Tilly (1993). Linear plasmids and chromosomes in bacteria. Molecular Microbiology 10(5), 917–922.
- Hjerde, E., M. S. Lorentzen, M. T. Holden, K. Seeger, S. Paulsen, N. Bason, C. Churcher, D. Harris, H. Norbertczak, M. A. Quail, , S. Sanders, S. Thurston, J. Parkhil, N. P. Willassen, and N. R. Thomson (2008). The genome sequence of the fish pathogen *Aliivibrio salmonicida* strain LFI1238 shows extensive evidence of gene decay. *BMC Genomics* 9(1), 616.

- Holden, M. T., R. W. Titball, S. J. Peacock, A. M. Cerdeño-Tárraga, T. Atkins, L. C. Crossman, T. Pitt, C. Churcher, K. Mungall, S. D. Bentley, M. Sebaihia, N. Thomson, N. Bason, I. R. Beacham, K. Brooks, K. A. Brown, N. F. Brown, G. L. Challis, I. Cherevach, T. Chillingworth, A. Cronin, B. Crossett, P. Davis, D. DeShazer, T. Feltwell, A. Fraser, Z. Hance, H. Hauser, S. Holroyd, K. Jagels, K. E. Keith, M. Maddison, S. Moule, C. Price, M. A. Quail, E. Rabbinowitsch, K. Rutherford, M. Sanders, M. Simmonds, S. Songsivilai, K. Stevens, S. Tumapa, M. Vesaratchavest, S. Whitehead, C. Yeats, B. G Barrell, P. C. F. Oyston and J. Parkhill (2004). Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proceedings of the National Academy of Sciences U.S.A. 101* (39), 14240–14245.
- Hornung, C., A. Poehlein, F. S. Haack, M. Schmidt, K. Dierking, A. Pohlen, H. Schulenburg, M. Blokesch, L. Plener, K. Jung, A. Bonge, I. Krohn-Molt, C. Utpatel, G. Timmermann, E. Spieck, A. Pommerening-Röser, E. Bode, H. B. Bode, R. Daniel, C. Schmeisser, W. R. Streit (2013). The *Janthinobacterium* sp. HH01 genome encodes a homologue of the V. cholerae CqsA and L. pneumophila LqsA autoinducer synthases. PloS One 8(2), e55045.
- Hosmer Jr, D. W., S. Lemeshow, and R. X. Sturdivant (2013). Applied logistic regression, Third Edition. Wiley Series in Probability and Statistics. Wiley-Blackwell.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology 24(6), 417.
- Husnik, F., N. Nikoh, R. Koga, L. Ross, R. P. Duncan, M. Fujie, M. Tanaka, N. Satoh, D. Bachtrog, A. C. Wilson, C. D. von Dohlen, T. Fukatsu, J. P. McCutcheon (2013). Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* 153(7), 1567–1578.
- Hyrien, O., A. Rappailles, G. Guilbaud, A. Baker, C.-L. Chen, A. Goldar, N. Petryk, M. Kahli, E. Ma, Y. d'Aubenton Carafa, B. Audit, C. Thermes and A. Arneodo (2013). From simple bacterial and archaeal replicons to replication N/U-domains. *Journal of Molecular Biology* 425(23), 4673–4689.
- Itaya, M. and T. Tanaka (1997). Experimental surgery to create subgenomes of Bacillus subtilis 168. Proceedings of the National Academy of Sciences U.S.A. 94(10), 5378– 5382.
- Izenman, A. (2008). Modern multivariate statistical techniques : regression. Classification, and Manifold Learning Series, Springer Texts in Statistics. Springer-Verlag, New York Inc.
- Jacob, F., S. Brenner, and F. Cuzin (1963). On the regulation of DNA replication in bacteria. Cold Spring Harbor Symposia on Quantitative Biology 28, 329–348.
- Jacomy, M., S. Heymann, T. Venturini, and M. Bastian (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* 9(6), e98679.

- Jain, A. and P. Srivastava (2013). Broad host range plasmids. FEMS Microbiology Letters 348(2), 87–96.
- Jensen, L. B., L. Garcia-Migura, A. J. S. Valenzuela, M. Løhr, H. Hasman, and F. M. Aarestrup (2010). A classification system for plasmids from enterococci and other Gram-positive bacteria. *Journal of Microbiological Methods* 80(1), 25–43.
- Jha, J. K., J. H. Baek, T. Venkova-Canova, and D. K. Chattoraj (2012). Chromosome dynamics in multichromosome bacteria. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1819(7), 826–829.
- Johnson, R. C., L. M. Johnson, J. W. Schmidt, and J. F. Gardner (2005). Major nucleoid proteins in the structure and function of the *Escherichia coli* chromosome. In N. P. Higgins (Ed.), *The Bacterial Chromosome*. ASM Press, Washington, DC. Pp. 65-132.
- Jumas-Bilak, E., S. Michaux-Charachon, G. Bourg, D. O'Callaghan, and M. Ramuz (1998). Differences in chromosome number and genome rearrangements in the genus Brucella. Molecular microbiology 27(1), 99–106.
- Kailing, K., H.-P. Kriegel, and P. Kröger (2004). Density-connected subspace clustering for high-dimensional data. In *Proceedings of SIAM International Conference on Data Mining (SDM'04), Orlando, FL.* Pp. 246–257.
- Kanehisa, M., S. Goto, Y. Sato, M. Furumichi, and M. Tanabe (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Re*search 40 (Database issue), D109–D114.
- Kaneko, T., Y. Nakamura, S. Sato, E. Asamizu, T. Kato, S. Sasamoto, A. Watanabe,
 K. Idesawa, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa,
 M. Kohara, M. Matsumoto, A. Matsuno, Y. Mochizuki, S. Nakayama, N. Nakazaki,
 S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada, and S. Tabata (2000). Complete
 genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Research* 7(6), 331–338.
- Kaneko, T., Y. Nakamura, C. P. Wolk, T. Kuritz, S. Sasamoto, A. Watanabe, M. Iriguchi,
 A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, M. Kohara, M. Matsumoto,
 A. Matsuno, A. Muraki, N. Nakazaki, S. Shimpo, M. Sugimoto, M. Takazawa,
 M. Yamada, M. Yasuda and S. Tabata (2001). Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. DNA Research 8(5), 205–213.
- Katayama, T., S. Ozaki, K. Keyamura, and K. Fujimitsu (2010). Regulation of the replication cycle : conserved and diverse regulatory systems for DnaA and oriC. Nature Reviews Microbiology 8(3), 163–170.
- Kawamukai, M., H. Matsuda, W. Fujii, R. Utsumi, and T. Komano (1989). Nucleotide sequences of *fic* and *fic-1* genes involved in cell filamentation induced by cyclic AMP in *Escherichia coli*. Journal of Bacteriology 171(8), 4525–4529.

- Kelly, W. J., S. C. Leahy, E. Altermann, C. J. Yeoman, J. C. Dunne, Z. Kong, D. M. Pacheco, D. Li, S. J. Noel, C. D. Moon, A. L. Cookson, and G. T. Attwood (2010). The glycobiome of the rumen bacterium *Butyrivibrio proteoclasticus* B316(T) highlights adaptation to a polysaccharide-rich environment. *PloS One* 5(8), e11942.
- Khan, S. A. (2005). Plasmid rolling-circle replication : highlights of two decades of research. *Plasmid* 53(2), 126–136.
- Kiełbasa, S. M., R. Wan, K. Sato, P. Horton, and M. C. Frith (2011). Adaptive seeds tame genomic sequence comparison. *Genome Research* 21(3), 487–493.
- Kiley, P. J. and S. Kaplan (1988). Molecular genetics of photosynthetic membrane biosynthesis in *Rhodobacter sphaeroides*. *Microbiological Reviews* 52(1), 50.
- King, G. and L. Zeng (2001). Logistic regression in rare events data. Political Analysis 9(2), 137–163.
- Kinscherf, T. G. and D. K. Willis (2002). Global regulation by gidA in Pseudomonas syringae. Journal of Bacteriology 184(8), 2281–2286.
- Kirkup, B. C., L. Chang, S. Chang, D. Gevers, and M. F. Polz (2010). Vibrio chromosomes share common history. BMC Microbiology 10(1), 137.
- Kiss, H., D. Cleland, A. Lapidus, S. Lucas, T. G. Del Rio, M. Nolan, H. Tice, C. Han, L. Goodwin, S. Pitluck, K. Liolios, N. Ivanova, K. Mavromatis, G. Ovchinnikova, A. Pati, A. Chen, K. Palaniappan, M. Land, L. Hauser, Y.-J. Chang, C. D. Jeffries, M. Lu, T. Brettin, J. C. Detter, M. Göker, B. J. Tindall, B. Beck, T. R. McDermott, T. Woyke, J. Bristow, J. A. Eisen, V. Markowitz, P. Hugenholtz, N. C. Kyrpides, H.-P. Klenk, and J.-F. Cheng (2010). Complete genome sequence of '*Thermobaculum terrenum*' type strain (YNP1). Standards in Genomic Sciences 3(2), 153–162.
- Kobayashi, I. (2004). Genetic addiction : a principle of gene symbiosis in a genome. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, ASM Press, Washington, DC. Pp. 105–114.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. Biological Cybernetics 43(1), 59–69.
- Kolstø, A.-B. (1997). Dynamic bacterial genome organization. Molecular microbiology 24(2), 241–248.
- Komatsu, H., Y. Imura, A. Ohori, Y. Nagata, and M. Tsuda (2003). Distribution and organization of auxotrophic genes on the multichromosomal genome of *Burkholderia multivorans* ATCC 17616. *Journal of Bacteriology* 185(11), 3333–3343.
- Kono, N., K. Arakawa, and M. Tomita (2011). Comprehensive prediction of chromosome dimer resolution sites in bacterial genomes. BMC Genomics 12(1), 19.
- Kono, N., K. Arakawa, and M. Tomita (2012). Validation of bacterial replication termination models using simulation of genomic mutations. *PloS One* 7(4), e34526.

- Koonin, E. V. (2000). How many genes can make a cell: The minimal-gene-set concept. Annual Review of Genomics and Human Genetics 1(1), 99–116.
- Koonin, E. V. (2003). The clusters of orthologous groups (COGs) database : phylogenetic classification of proteins from complete genomes. In J. McEntyre and J. Ostell (Eds.), *The NCBI Handbook [Internet]*, Chapter 22. National Center for Biotechnology Information (US).
- Korf, I., M. Yandell, and J. Bedell (2003). An Essential Guide to the Basic Local Alignment Search Tool : BLAST. O'Reilly Media, Inc.
- Kovács, F., C. Legány, and A. Babos (2005). Cluster validity measurement techniques. In Proceedings of the 6th International Symposium of Hungarian Researchers on Computational Intelligence, Budapest, Nov. 2005, 18-19.
- Krone, S. M., R. Lu, R. Fox, H. Suzuki, and E. M. Top (2007). Modelling the spatial dynamics of plasmid transfer and persistence. *Microbiology* 153(8), 2803–2816.
- Krüger, R. and S. A. Rakowski (2004). Participating elements in the replication of iteron containing plasmids. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid biology*, ASM Press, Washington, DC. Pp. 25–62.
- Lackner, G., N. Moebius, L. Partida-Martinez, and C. Hertweck (2011). Complete genome sequence of *Burkholderia rhizoxinica*, an endosymbiont of *Rhizopus microsporus*. *Journal of Bacteriology* 193(3), 783–784.
- Lan, R. and P. R. Reeves (1996). Gene transfer is a major factor in bacterial evolution. Molecular Biology and Evolution 13(1), 47–55.
- Lan, R., P. R. Reeves, and S. Octavia (2009). Population structure, origins and evolution of major Salmonella enterica clones. Infection, Genetics and Evolution 9(5), 996– 1005.
- Lancichinetti, A. and S. Fortunato (2009). Community detection algorithms : a comparative analysis. *Physical Review E* 80(5), 056117.
- Landeta, C., A. Dávalos, M. A. Cevallos, O. Geiger, S. Brom, and D. Romero (2011). Plasmids with a chromosome-like role in rhizobia. *Journal of Bacteriology* 193(6), 1317–1326.
- Lanoil, B. D., L. M. Ciuffetti, and S. J. Giovannoni (1996). The marine bacterium *Pseudoalteromonas haloplanktis* has a complex genome structure composed of two separate genetic units. *Genome Research* 6(12), 1160–1169.
- Larose, D. T. (2006). Data Mining Methods and Models. Wiley-IEEE Press.
- Lassalle, F. (2013). Les génomes bactériens, une histoire de transferts de gènes, de recombinaison et de cladogénèse. Thèse de doctorat en Physiologie et biologie des organismes populations interactions, Université de Lyon.

- Lau, S. Y. and H. I. Zgurskaya (2005). Cell division defects in *Escherichia coli* deficient in the multidrug efflux transporter AcrEF-TolC. *Journal of Bacteriology* 187(22), 7815–7825.
- Lawley, T., B. M. Wilkins, and L. S. Frost (2004). Bacterial conjugation in Gramnegative bacteria. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, ASM press, Washington, DC. Pp. 203–226.
- Lederberg, J. (1952). Cell genetics and hereditary symbiosis. *Physiological Review* 32(4), 403–430.
- Leplae, R., G. Lima-Mendez, and A. Toussaint (2010). ACLAME : a CLAssification of Mobile genetic Elements, update 2010. Nucleic Acids Research 38 (Database issue), D57–D61.
- Leroux, M., F. Jia, and G. Szatmari (2011). Characterization of the Streptococcus suis XerS recombinase and its unconventional cleavage of the difSL site. FEMS Microbiology Letters 324(2), 135–141.
- Li, X., M. Wu, C.-K. Kwoh, and S.-K. Ng (2010). Computational approaches for detecting protein complexes from protein interaction networks : a survey. BMC Genomics 11 (Suppl 1), S3.
- Li, Z., A. L. Garner, C. Gloeckner, K. D. Janda, and C. K. Carlow (2011). Targeting the Wolbachia cell division protein FtsZ as a new approach for antifilarial therapy. *PLoS Neglected Tropical Diseases* 5(11), e1411.
- Lili, L. N., N. F. Britton, and E. J. Feil (2010). The persistence of parasitic plasmids. Genetics 177(1), 399–405.
- Lima-Mendez, G., J. Van Helden, A. Toussaint, and R. Leplae (2008). Reticulate representation of evolutionary and functional relationships between phage genomes. *Mole*cular Biology and Evolution 25(4), 762–777.
- Livny, J., Y. Yamaichi, and M. K. Waldor (2007). Distribution of centromere-like parS sites in bacteria : insights from comparative genomics. *Journal of Bacteriology* 189(23), 8693–8703.
- Loftie-Eaton, W. and D. E. Rawlings (2012). Diversity, biology and evolution of IncQ-family plasmids. *Plasmid* 67(1), 15–34.
- Lopes, A., J. Amarir-Bouhram, G. Faure, M.-A. Petit, and R. Guerois (2010). Detection of novel recombinases in bacteriophage genomes unveils Rad52, Rad51 and Gp2.5 remote homologs. *Nucleic Acids Research* 38(12), 3952–3962.
- López-Guerrero, M. G., E. Ormeño-Orrillo, J. L. Acosta, A. Mendoza-Vargas, M. A. Rogel, M. A. Ramírez, M. Rosenblueth, J. Martínez-Romero, and E. Martínez-Romero (2012). Rhizobial extrachromosomal replicon variability, stability and expression in natural niches. *Plasmid* 68(3), 149–158.

- López-Madrigal, S., A. Latorre, M. Porcar, A. Moya, and R. Gil (2011). Complete genome sequence of "*Candidatus* Tremblaya princeps" strain PCVAL, an intriguing translational machine below the living-cell status. *Journal of Bacteriology* 193(19), 5587–5588.
- Mackenzie, C., M. Choudhary, F. W. Larimer, P. F. Predki, S. Stilwagen, J. P. Armitage, R. D. Barber, T. J. Donohue, J. P. Hosler, J. E. Newman, J. P. Shapleigh, R. E. Sockett, J. Zeilstra-Ryalls and S. Kaplan(2001). The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4. 1. *Photosynthesis Research* 70(1), 19–41.
- Mackenzie, C., J. M. Eraso, M. Choudhary, J. H. Roh, X. Zeng, P. Bruscella, A. Puskás, and S. Kaplan (2007). Postgenomic adventures with *Rhodobacter sphaeroides*. Annual *Review of Microbiology* 61, 283–307.
- Mackenzie, C., S. Kaplan, and M. Choudhary (2004). Multiple chromosomes. In Microbial evolution : gene establishment, survival, and exchange, ASM press, Washington, DC. Pp. 82–101.
- MacLellan, S. R., C. D. Sibley, and T. M. Finan (2004). Second chromosomes and megaplasmids in bacteria. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*. ASM press, Washington, DC. Pp. 529–542.
- Maida, I., M. Fondi, V. Orlandini, G. Emiliani, M. C. Papaleo, E. Perrin, and R. Fani (2014). Origin, duplication and reshuffling of plasmid genes : Insights from *Burkholderia vietnamiensis* G4 genome. *Genomics* 103(2-3), 229–238.
- Maj, A., L. Dziewit, J. Czarnecki, M. Włodarczyk, J. Baj, G. Skrzypczyk, D. Giersz, and D. Bartosik (2013). Plasmids of carotenoid-producing *Paracoccus* spp. (Alphaproteobacteria) structure, diversity and evolution. *PloS One* 8(11), e80258.
- Marchler-Bauer, A., J. B. Anderson, M. K. Derbyshire, C. DeWeese-Scott, N. R. Gonzales, M. Gwadz, L. Hao, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, G. H. Marchler, M. Mullokandov, B. A. Shoemaker, V. Simonyan, J. S Song, P. A. Thiessen, R. A. Yamashita, J. J. Yin, Zhang, D. and S. H. Bryant (2007). CDD : a conserved domain database for interactive domain family analysis. *Nucleic Acids Research* 35 (Database issue), D237–D240.
- Martinez, R. J., D. Bruce, C. Detter, L. A. Goodwin, J. Han, C. S. Han, B. Held, M. L. Land, N. Mikhailova, M. Nolan, L. Pennacchioc, S. Pitluckc, R. Tapiab, T. Woykec and P. A. Sobecky (2012). Complete genome sequence of *Rahnella* sp. Strain Y9602, a gammaproteobacterium isolate from metal-and radionuclide-contaminated soil. *Journal of Bacteriology* 194(8), 2113–2114.
- McGeoch, A. T. and S. D. Bell (2008). Extra-chromosomal elements and the evolution of cellular DNA replication machineries. *Nature Reviews Molecular Cell Biology* 9(7), 569–574.
- Medema, M. H., A. Trefzer, A. Kovalchuk, M. van den Berg, U. Müller, W. Heijne, L. Wu, M. T. Alam, C. M. Ronning, W. C. Nierman, et al. (2010). The sequence of

a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. *Genome Biology and Evolution 2*, 212–224.

- Médigue, C., E. Krin, G. Pascal, V. Barbe, A. Bernsel, P. N. Bertin, F. Cheung, S. Cruveiller, S. D'Amico, A. Duilio, G. Fang, G. Feller, C. Ho, S. Mangenot, G. Marino, J. Nilsson, E. Parrilli, E. P. C. Rocha, Z. Rouy, A. Sekowska, M. L. Tutino, D. Vallenet, G. von Heijne and A. Danchin (2005). Coping with cold : the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Research* 15(10), 1325–1335.
- Mendell, J. E., K. D. Clements, J. H. Choat, and E. R. Angert (2008). Extreme polyploidy in a large bacterium. *Proceedings of the National Academy of Sciences* U.S.A. 105(18), 6730–6734.
- Messer, W. (2002). The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode to initiate DNA replication. *FEMS Microbiology Reviews* 26(4), 355–374.
- Michaux, S., J. Paillisson, M. Carles-Nurit, G. Bourg, A. Allardet-Servent, and M. Ramuz (1993). Presence of two independent chromosomes in the *Brucella melitensis* 16M genome. *Journal of Bacteriology* 175(3), 701–705.
- Mierzejewska, J. and G. Jagura-Burdzy (2012). Prokaryotic ParA-ParB-parS system links bacterial chromosome segregation with the cell cycle. *Plasmid* 67(1), 1–14.
- Miller, R. V. (2004). Bacteriophage-mediated transduction : An engine for change and evolution. In R. V. Miller and M. J. Day (Eds.), *Microbial Evolution : Gene Establishment, Survival, and Exchange*, ASM Press, Washington, DC. Pp. 144–157.
- Miller, R. V. and M. J. Day (2004). *Microbial evolution : gene establishment, survival, and exchange.* ASM press, Washington, DC.
- Mochizuki, A., K. Yahara, I. Kobayashi, and Y. Iwasa (2006). Genetic addiction : selfish gene's strategy for symbiosis in the genome. *Genetics* 172(2), 1309–1323.
- Moran, M., R. Belas, M. Schell, J. Gonzalez, F. Sun, S. Sun, B. Binder, J. Edmonds, W. Ye, B. Orcutt, E. C. Howard, C. Meile, W. Palefsky, A. Goesmann, Q. Ren, I. Paulsen, L. E. Ulrich, L. S. Thompson, E. Saunders and A. Buchan (2007). Ecological genomics of marine roseobacters. *Applied and environmental Microbiology* 73(14), 4559–4569.
- Moran, M. A., A. Buchan, J. M. González, J. F. Heidelberg, W. B. Whitman, R. P. Kiene, J. R. Henriksen, G. M. King, R. Belas, C. Fuqua, L. Brinkac, M. R. Lewis, S. Johri, B. Weaver, G. Pal, J. A. Eisen, E. Rahe, W. M. Sheldon, W. Ye, T. R. Miller, J. Carlton, D. A. Rasko, I. T. Paulsen, Q. Ren, S. C. Daugherty, R. T. Deboy, R. J. Dodson, A. S. Burkin, R. Madupu, W. C. Nelson, S. A. Sullivan, M. J. Rosovitz, D. H. Haft, J. Selengut, and N. Ward (2004). Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* 432(7019), 910–913.
- Moreno, E. (1998). Genome evolution within the alpha Proteobacteria : why do some bacteria not possess plasmids and others exhibit more than one different chromosome ? *FEMS Microbiology Reviews* 22(4), 255–275.

- Morita, R. Y. (1988). Bioavailability of energy and its relationship to growth and starvation survival in nature. *Canadian Journal of Microbiology* 34(4), 436–441.
- Morrow, J. D. and V. S. Cooper (2012). Evolutionary effects of translocations in bacterial genomes. *Genome Biology and Evolution* 4(12), 1256–1262.
- Mott, M. L. and J. M. Berger (2007). DNA replication initiation : mechanisms and regulation in bacteria. *Nature Reviews Microbiology* 5(5), 343–354.
- Nagata, Y., M. Matsuda, H. Komatsu, Y. Imura, H. Sawada, Y. Ohtsubo, and M. Tsuda (2005). Organization and localization of the *dnaA* and *dnaK* gene regions on the multichromosomal genome of *Burkholderia multivorans* ATCC 17616. *Journal of Bioscience* and *Bioengineering* 99(6), 603–610.
- Nagata, Y., Y. Ohtsubo, R. Endo, N. Ichikawa, A. Ankai, A. Oguchi, S. Fukui, N. Fujita, and M. Tsuda (2010). Complete genome sequence of the representative γhexachlorocyclohexane-degrading bacterium Sphingobium japonicum UT26. Journal of Bacteriology 192(21), 5852–5853.
- Naïm, P., P.-H. Wuillemin, P. Leray, O. Pourret, and A. Becker (2011). Réseaux bayésiens. Editions Eyrolles.
- Nascimento, A. L. T. O., A. I. Ko, E. A. L. Martins, C. B. Monteiro-Vitorello, P. L. Ho P. Ho, D. A. Haake, S. Verjovski-Almeida, R. A. Hartskeerl, M. V. Marques, M. C. Oliveira, C. F. M. Menck, L. C. C.Leite, H. Carrer, L. L. Coutinho, W. M. Degrave, O. A. Dellagostin, H. El-Dorry, E. S.Ferro, M. I. T. Ferro, L. R. Furlan, M. Gamberini, E. A. Giglioti, A. Góes-Neto, G. H. Goldman, M. H. S. Goldman, R. Harakava, S. M. B. Jerônimo, I. L. M. Junqueira-de Azevedo, E. T. Kimura, E. E. Kuramae, E. G. M. Lemos, M. V. F. Lemos, C. L. Marino, L. R. Nunes, R. C. de Oliveira, G. G. Pereira, M. S. Reis, A. Schriefer, W. J. Siqueira, P. Sommer, S. M. Tsai, A. J. G. Simpson, J. A. Ferro, L. E. A. Camargo, J. P. Kitajima, J. Setubal, and M. A. V. Sluys (2004). Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *Journal of Bacteriology* 186(7), 2164–2172.
- Needleman, S. B. and C. D. Wunsch (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* 48(3), 443–453.
- Nepusz, T., H. Yu, and A. Paccanaro (2012). Detecting overlapping protein complexes in protein-protein interaction networks. *Nature Methods* 9(5), 471–472.
- Nierman, W. C., D. DeShazer, H. S. Kim, H. Tettelin, K. E. Nelson, T. Feldblyum, R. L. Ulrich, C. M. Ronning, L. M. Brinkac, S. C. Daugherty, T. D. Davidsen, R. T. Deboy, G. Dimitrov, R. J. Dodson, A. S. Durkin, M. L. Gwinn, D. H. Hafft, H. Khouri, J. F. Kolonay, R. Madupu, Y. Mohammoud, W. C. Nelson, D. Radune, C. M. Romero, S. Sarria, J. Selengut, C. Shamblin, S. A. Sullivan, O. White, Y. Yu, N. Zafar, L. Zhou, and C. Fraser (2004). Structural flexibility in the *Burkholderia mallei* genome. *Proceedings of the National Academy of Sciences U.S.A. 101*(39), 14246–14251.

- Ochman, H. (2002). Bacterial evolution : chromosome arithmetic and geometry. Current biology 12(12), R427–R428.
- O'Donnell, M., L. Langston, and B. Stillman (2013). Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harbor perspectives in biology* 5(7), a010108.
- Ogasawara, N. and H. Yoshikawa (1992). Genes and their organization in the replication origin region of the bacterial chromosome. *Molecular Microbiology* 6(5), 629–634.
- Omelchenko, M. V., Y. I. Wolf, E. K. Gaidamakova, V. Y. Matrosova, A. Vasilenko, M. Zhai, M. J. Daly, E. V. Koonin, and K. S. Makarova (2005). Comparative genomics of *Thermus thermophilus* and *Deinococcus radiodurans* : divergent routes of adaptation to thermophily and radiation resistance. *BMC Evolutionary Biology* 5(1), 57.
- O'Rourke, S., A. Wietzorrek, K. Fowler, C. Corre, G. L. Challis, and K. F. Chater (2009). Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in *Streptomyces coelicolor*. *Molecular Microbiology* 71(3), 763–778.
- O'Sullivan, J. M. (2011). Chromosome organization in simple and complex unicellular organisms. *Current issues in Molecular Biology* 13(2), 37–42.
- Passot, F. M., V. Calderon, G. Fichant, D. Lane, and F. Pasta (2012). Centromere binding and evolution of chromosomal partition systems in the burkholderiales. *Journal* of Bacteriology 194(13), 3426–3436.
- Pati, A., K. LaButti, R. Pukall, M. Nolan, T. Glavina Del Rio, H. Tice, J.-F. Cheng, S. Lucas, F. Chen, A. Copeland, N. Ivanova, K. Mavromatis, N. Mikhailova, S. Pitluck, D. Bruce, L. Goodwin, M. Land, L. Hauser, Y. -J. Chang, C. D. Jeffries, A. Chen, K. Palaniappan, P. Chain, T. Brettin, J. Sikorski, M. Rohde, M. Göker, J. Bristow, J. A. Eisen, V. Markowitz, P. Hugenholtz, N. C. Kyrpides, H. -P. Klenk, and A. Lapidus (2010). Complete genome sequence of Sphaerobacter thermophilus type strain (S 6022). Standards in Genomic Sciences 2(1), 49–56.
- Paulsen, I. T., R. Seshadri, K. E. Nelson, J. A. Eisen, J. F. Heidelberg, T. D. Read, R. J. Dodson, L. Umayam, L. M. Brinkac, M. J. Beanan, S. C. Daugherty, R. T. Deboy, A. S. Durkin, J. F. Kolonay, R. Madupu, W. C. Nelson, B. Ayodeji, M. Kraul, J. Shetty, J. Malek, S. E. Van Aken, S. Riedmuller, Herve Tettelin, S. R. Gill, O. White, S. L. Salzberg, D. L. Hoover, L. E. Lindler, S. M. Halling, S. M. Boyle, and C. M. Fraser (2002). The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proceedings of the National Academy of Sciences U.S.A. 99*(20), 13148–13153.
- Pérals, K., F. Cornet, Y. Merlet, I. Delon, and J.-M. Louarn (2000). Functional polarization of the *Escherichia coli* chromosome terminus : the *dif* site acts in chromosome dimer resolution only when located between long stretches of opposite polarity. *Molecular Microbiology* 36(1), 33–43.

- Perry, J. J. and J. T. Staley (1997). *Microbiology : dynamics and diversity*. Saunders College Publishing, Harcourt Brace College Publishers.
- Petersen, J. (2011). Phylogeny and compatibility : plasmid classification in the genomics era. Archives of Microbiology 193(5), 313–321.
- Petersen, J., H. Brinkmann, M. Berger, T. Brinkhoff, O. Päuker, and S. Pradella (2011). Origin and evolution of a novel DnaA-like plasmid replication type in Rhodobacterales. *Molecular Biology and Evolution* 28(3), 1229–1240.
- Petersen, J., O. Frank, M. Göker, and S. Pradella (2013, February). Extrachromosomal, extraordinary and essential-the plasmids of the Roseobacter clade. *Applied microbio*logy and biotechnology, 2805–2815.
- Picardeau, M., D. M. Bulach, C. Bouchier, R. L. Zuerner, N. Zidane, P. J. Wilson, S. Creno, E. S. Kuczek, S. Bommezzadri, J. C. Davis, A. McGrath, M. J. Johnson, C. Boursaux-Eude, T. Seemann, Z. Rouy, R. L. Coppel, J. I. Rood, A. Lajus, J. K. Davies, C. Médigue, and B. Adler (2008). Genome sequence of the saprophyte *Leptospira biflexa* provides insights into the evolution of *Leptospira* and the pathogenesis of leptospirosis. *PloS One* 3(2), e1607.
- Pinto, U. M., K. M. Pappas, and S. C. Winans (2012). The ABCs of plasmid replication and segregation. *Nature Reviews Microbiology* 10(11), 755–765.
- Pohlmann, A., W. F. Fricke, F. Reinecke, B. Kusian, H. Liesegang, R. Cramm, T. Eitinger, C. Ewering, M. Pötter, E. Schwartz, A. Strittmatter, I. Voß, G. Gottschalk, A. Steinbüchel, B. Friedrich, and B. Bowien (2006). Genome sequence of the bioplastic-producing "Knallgas" bacterium *Ralstonia eutropha* H16. *Nature Biotechnology* 24 (10), 1257–1262.
- Pruitt, K. D., T. Tatusova, and D. R. Maglott (2007). NCBI reference sequences (Ref-Seq) : a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research* 35(Database issue), D61–D65.
- Purushe, J., D. E. Fouts, M. Morrison, B. A. White, R. I. Mackie, P. M. Coutinho, B. Henrissat, and K. E. Nelson (2010). Comparative genome analysis of *Prevotella ruminicola* and *Prevotella bryantii* : insights into their environmental niche. *Microbial Ecology* 60(4), 721–729.
- Qiao, J., L. Chen, Y. Li, J. Wang, W. Zhang, and S. Chen (2012). Whole-genome sequence of Nocardiopsis alba strain ATCC BAA-2165, associated with honeybees. *Journal of Bacteriology* 194 (22), 6358–6359.
- Qin, Q.-L., Y. Li, Y.-J. Zhang, Z.-M. Zhou, W.-X. Zhang, X.-L. Chen, X.-Y. Zhang, B.-C. Zhou, L. Wang, and Y.-Z. Zhang (2011). Comparative genomics reveals a deep-sea sediment-adapted life style of *Pseudoalteromonas* sp. SM9913. *The ISME Journal* 5(2), 274–284.
- Rajewska, M., K. Wegrzyn, and I. Konieczny (2012). AT-rich region and repeated sequences-the essential elements of replication origins of bacterial replicons. *FEMS Microbiology Reviews* 36(2), 408–434.

- Rakhlin, A. and A. Caponnetto (2007). Stability of k-means clustering. Advances in Neural Information Processing Systems 19, 1121.
- Ramírez-Bahena, M. H., L. Vial, F. Lassalle, B. Diel, D. Chapulliot, V. Daubin, X. Nesme, and D. Muller (2014). Single acquisition of protelomerase gave rise to speciation of a large and diverse clade within the Agrobacterium/Rhizobium supercluster characterized by the presence of a linear chromid. Molecular phylogenetics and evolution 73, 202–207.
- Rasmussen, T., R. B. Jensen, and O. Skovgaard (2007). The two chromosomes of Vibrio cholerae are initiated at different time points in the cell cycle. The EMBO Journal 26(13), 3124–3131.
- Ren, S.-X., G. Fu, X.-G. Jiang, R. Zeng, Y.-G. Miao, H. Xu, Y.-X. Zhang, H. Xiong, G. Lu, L.-F. Lu, H. -Q. Jiang, J. Jia, Y. -F. Tu, J. -X. Jiang, W. -Y. Gu, Y. -Q. Zhang, Z. Cai, H. -Hui Sheng, H. -Feng Yin, Y. Zhang, G. -F. Zhu, Ma Wan, H. -L. Huang, Z. Qian, S. -Y. Wang, W. Ma, Z. -J. Yao, Y. Shen, B. -Q. Qiang, Q. C. Xia, X. -K. Guo, A. Danchin, I. Saint Girons, R. L. Somerville, Y. -M. Wen, M. -H. Shi, Z. Chen, J. -G. Xu, and G. -P. Zhao (2003). Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. Nature 422(6934), 888–893.
- Rendón, E., I. Abundez, A. Arizmendi, and E. M. Quiroz (2011). Internal versus External cluster validation indexes. *International Journal of Computers and Communica*tions 5(1), 27–34.
- Renwick, J. (1971). The Rhesus syntenic group in man. Nature 234 (5330), 475.
- Riley, M. A. and M. Lizotte-Waniewski (2009). Population genomics and the bacterial species concept. *Methods In Molecular Biology* 532(16), 367–377.
- Rimsky, S. and A. Travers (2011). Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins. *Current Opinion in Microbiologypinion in microbio*logy 14(2), 136–141.
- Roberts, A. P., M. Chandler, P. Courvalin, G. Guédon, P. Mullany, T. Pembroke, J. I. Rood, C. J. Smith, A. O. Summers, M. Tsuda, and D. E. Berg (2008). Revised nomenclature for transposable genetic elements. *Plasmid* 60(3), 167–173.
- Robinson, N. P. and S. D. Bell (2005). Origins of DNA replication in the three domains of life. *The FEBS Journal 272*(15), 3757–366.
- Rodley, P. D., U. Römling, and B. Tümmler (1995). A physical genome map of the Burkholderia cepacia type strain. Molecular Microbiology 17(1), 57–67.
- Rose, P. W., C. Bi, W. F. Bluhm, C. H. Christie, D. Dimitropoulos, S. Dutta, R. K. Green, D. S. Goodsell, A. Prlić, M. Quesada, G. B. Quinn, A. G. Ramos, J. D. Westbrook, J. Young, C. Zardecki, H. M. Berman, and P. E. Bourne (2013). The RCSB Protein Data Bank : new resources for research and education. *Nucleic Acids Research* 41 (Database issue), D475–D482.

- Rosenberg, A. and J. Hirschberg (2007). V-Measure : A conditional entropy-based external cluster evaluation measure. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), Number June, pp. 410–420.
- Rosvall, M. and C. T. Bergstrom (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* U.S.A. 105(4), 1118–1123.
- Rousseeuw, P. J. (1987). Silhouettes : a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20, 53–65.
- Rüping, S. (2004). A simple method for estimating conditional probabilities for SVMs. In A. Abecker and S. Weibelza hl (Eds.), *LWA*, Humbold-Universitat Berlin. Pp. 206–210.
- Saint-Dic, D., J. Kehrl, B. Frushour, and L. S. Kahng (2008). Excess SeqA leads to replication arrest and a cell division defect in Vibrio cholerae. Journal of Bacteriology 190(17), 5870–5878.
- Salanoubat, M., S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Arlat, A. Billault, P. Brottier, J. Camus, L. Cattolico, M. Chandler, N. Choisne, C. Claudel-Renard, S. Cunnac, N. Demange, C. Gaspin, M. Lavie, A. Moisan, C. Robert, W. Saurin, T. Schiex, P. Siguier, P. Thébault, M. Whalen, P. Wincker, M. Levy, J. Weissenbach, and C. A. Boucher (2002). Genome sequence of the plant pathogen *Ralstonia* solanacearum. Nature 415(6871), 497–502.
- Samson, R. Y. and S. D. Bell (2011). Cell cycles and cell division in the archaea. Current Opinion in Microbiology 14(3), 350–356.
- Sikorski, J., O. Chertkov, A. Lapidus, M. Nolan, S. Lucas, T. G. Del Rio, H. Tice, J.-F. Cheng, R. Tapia, C. Han, L. Goodwin, S. Pitluck, K. Liolios, N. Ivanova, K. Mavromatis, N. Mikhailova, A. Pati, A. Chen, K. Palaniappan, M. Land, L. Hauser, Y.-J. Chang, C. D. Jeffries, E. Brambilla, M. Yasawong, M. Rohde, R. Pukall, S. Spring, M. Göker, T. Woyke, J. Bristow, J. A. Eisen, V. Markowitz, P. Hugenholtz, N. C. Kyrpides, and H.-P. Klenk (2010). Complete genome sequence of *Ilyobacter polytropus* type strain (CuHbu1). *Standards in Genomic Sciences* 3(3), 304–314.
- Skarstad, K., E. Boye, and H. B. Steen (1986). Timing of initiation of chromosome replication in individual *Escherichia coli* cells. *The EMBO Journal* 5(7), 1711–1717.
- Slater, F., M. Bailey, A. Tett, and S. Turner (2008). Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Micriobiology Ecology* 66(1), 3–13.
- Slater, S. C., B. S. Goldman, B. Goodner, J. A. C. Setubal, S. K. Farrand, E. W. Nester, T. J. Burr, L. Banta, A. W. Dickerman, I. Paulsen, L. Otten, G. Suen, R. Welch, N. F. Almeida, F. Arnold, O. T. Burton, Z. Du, A. Ewing, E. Godsy, S. Heisel, K. L. Houmiel, J. Jhaveri, J. Lu, N. M. Miller, S. Norton, Q. Chen, W. Phoolcharoen, V. Ohlin, D. Ondrusek, N. Pride, S. L. Stricklin, J. Sun, C. Wheeler, L. Wilson,

H. Zhu, and D. W. Wood (2009, April). Genome sequences of three Agrobacterium biovars help elucidate the evolution of multichromosome genomes in bacteria. Journal of bacteriology 191(8), 2501–2511.

- Smillie, C., M. P. Garcillán-Barcia, M. V. Francia, E. P. C. Rocha, and F. De La Cruz (2010). Mobility of plasmids. *Microbiology and Molecular Biology Reviews* 74(3), 434–452.
- Smith, T. F. and M. S. Waterman (1981). Identification of common molecular subsequences. Journal of Molecular Biology 147(1), 195–197.
- Song, N., J. M. Joseph, G. B. Davis, and D. Durand (2008). Sequence similarity network reveals common ancestry of multidomain proteins. *PLoS Computational Biology* 4(5), e1000063.
- Song, N., R. D. Sedgewick, and D. Durand (2007). Domain architecture comparison for multidomain homology identification. *Journal of Computational Biology* 14(4), 496–516.
- Songsivilai, S. and T. Dharakul (2000). Multiple replicons constitute the 6.5-megabase genome of *Burkholderia pseudomallei*. Acta tropica 74, 169–179.
- Spies, M. and S. C. Kowalczykowski (2005). Homologous recombination by RecBCD and RecF pathways. In N. P. Higgins (Ed.), *The bacterial Chromosome*, ASM Press, Washington, DC. Pp. 389–403.
- Srivastava, P. and D. K. Chattoraj (2007). Selective chromosome amplification in Vibrio cholerae. Molecular Microbiology 66(4), 1016–1028.
- Stewart, P., P. A. Rosa, and K. Tilly (2004). Linear plasmids in bacteria : common origins, uncommon ends. In *Plasmid Biology*, ASM Press, Washington, DC. Pp. 291– 301.
- Stokke, C., T. Waldminghaus, and K. Skarstad (2011). Replication patterns and organization of replication forks in *Vibrio cholerae*. *Microbiology* 157(3), 695–708.
- Stouthamer, A. and S. Kooijman (1993). Why it pays for bacteria to delete disused DNA and to maintain megaplasmids. Antonie van Leeuwenhoek 63(1), 39–43.
- Summers, D. K. and D. J. Sherratt (1984). Multimerization of high copy number plasmids causes instability : ColE 1 encodes a determinant essential for plasmid monomerization and stability. *Cell 36*(4), 1097–1103.
- Sun, H., A. Lapidus, M. Nolan, S. Lucas, T. G. Del Rio, H. Tice, J.-F. Cheng, R. Tapia, C. Han, L. Goodwin, S. Pitluck, I. Pagani, N. Ivanova, K. Mavromatis, N. Mikhailova, A. Pati, A. Chen, K. Palaniappan, M. Land, L. Hauser, Y.-J. Chang, C. D. Jeffries, O. D. N. Djao, M. Rohde, J. Sikorski, M. Göker, T. Woyke, J. Bristow, J. A. Eisen, V. Markowitz, P. Hugenholtz, N. C. Kyrpides, and H.-P. Klenk (2010). Complete genome sequence of *Nocardiopsis dassonvillei* type strain (IMRU 509). *Standards in Genomic Sciences* 3(3), 325–336.

- Suwanto, A. and S. Kaplan (1989a). Physical and genetic mapping of the *Rhodobacter* sphaeroides 2.4. 1 genome : genome size, fragment identification, and gene localization. Journal of Bacteriology 171(11), 5840–5849.
- Suwanto, A. and S. Kaplan (1989b). Physical and genetic mapping of the *Rhodobacter* sphaeroides 2.4. 1 genome : presence of two unique circular chromosomes. Journal of Bacteriology 171(11), 5850–5859.
- Swingley, W. D., M. Chen, P. C. Cheung, A. L. Conrad, L. C. Dejesa, J. Hao, B. M. Honchak, L. E. Karbach, A. Kurdoglu, S. Lahiri, S. D. Mastrian, H. Miyashita, L. Page, P. Ramakrishna, S. Satoh, W. M. Sattley, Y. Shimada, H. L. Taylor, T. Tomo, T. Tsuchiya, Z. T. Wang, J. Raymond, M. Mimuro, R. E. Blankenship, and J. W. Touchman (2008). Niche adaptation and genome expansion in the chlorophyll *d*-producing cyanobacterium Acaryochloris marina. Proceedings of the National Academy of Sciences U.S.A. 105(6), 2005–2010.
- Tagomori, K., T. Iida, and T. Honda (2002). Comparison of genome structures of vibrios, bacteria possessing two chromosomes. *Journal of Bacteriology* 184(16), 4351–4358.
- Tenenbaum, J. B. (1998). Mapping a manifold of perceptual observations. In M. K. Michael I Jordan and S. A. Solla (Eds.), Proceeding NIPS '97 Proceedings of the 1997 conference on Advances in neural information processing systems 10, MIT Press Cambridge, MA. Pp. 682–688.
- Tenenbaum, J. B., V. De Silva, and J. C. Langford (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* 290 (5500), 2319–2323.
- Terrapon, N., J. Weiner, S. Grath, A. D. Moore, and E. Bornberg-Bauer (2014). Rapid similarity search of proteins using alignments of domain arrangements. *Bioinforma*tics 30(2), 274–281.
- Thanbichler, M. (2010). Synchronization of chromosome dynamics and cell division in bacteria. *Cold Spring Harbor perspectives in biology* 2(1), a000331.
- Thomas, C. (2004). Evolution and population genetics of bacterial plasmids. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid Biology*, ASM press, Washington, DC. Pp. 509–528.
- Thompson, F. L., T. Iida, and J. Swings (2004). Biodiversity of vibrios. *Microbiology* and *Molecular Biology Reviews* 68(3), 403–431.
- Tibshirani, R. and G. Walther (2005). Cluster Validation by Prediction Strength. Journal of Computational and Graphical Statistics 14(3), 511–528.
- Tichi, M. A. and F. R. Tabita (2001). Interactive control of *Rhodobacter capsula*tus redox-balancing systems during phototrophic metabolism. Journal of Bacteriology 183(21), 6344–6354.
- Tobiason, D. M. and H. S. Seifert (2006). The obligate human pathogen, *Neisseria gonorrhoeae*, is polyploid. *PLoS Biology* 4(6), e185.

- Toro, E. and L. Shapiro (2010). Bacterial chromosome organization and segregation. Cold Spring Harbor perspectives in biology 2(2), a000349.
- Toukdarian, A. (2004). Plasmid strategies for broad-host-range replication in gramnegative bacteria. In B. E. Funnel and G. J. Phillips (Eds.), *Plasmid biology*, ASM Press, Washington, DC. Pp. 259–270.
- Touzain, F., M.-A. Petit, S. Schbath, and M. El Karoui (2011). DNA motifs that sculpt the bacterial chromosome. *Nature Reviews Microbiology* 9(1), 15–26.
- Trucksis, M., J. Michalski, Y. K. Deng, and J. B. Kaper (1998). The Vibrio cholerae genome contains two unique circular chromosomes. Proceedings of the National Academy of Sciences U.S.A. 95(24), 14464–14469.
- Val, M.-E., S. P. Kennedy, M. El Karoui, L. Bonné, F. Chevalier, and F.-X. Barre (2008). FtsK-dependent dimer resolution on multiple chromosomes in the pathogen *Vibrio cholerae*. *PLoS Genetics* 4(9), e1000201.
- Val, M.-E., S. P. Kennedy, A. J. Soler-Bistué, V. Barbe, C. Bouchier, M. Ducos-Galand, O. Skovgaard, and D. Mazel (2014). Fuse or die : how to survive the loss of Dam in *Vibrio cholerae. Molecular Microbiology 91*(4), 665–678.
- van Dongen, S. M. (2000). *Graph clustering by flow simulation*. Ph.D. thesis, University of Utrecht.
- Van Houdt, R., R. Leplae, G. Lima-Mendez, M. Mergeay, and A. Toussaint (2012). Towards a more accurate annotation of tyrosine-based site-specific recombinases in bacterial genomes. *Mobile DNA* 3(1), 6.
- Venkova-Canova, T. C. D. K. (2011). Transition from a plasmid to a chromosomal mode of replication entails additional regulators. *Proceedings of the National Academy of Sciences U.S.A. 108*(15), 6199–6204.
- Vezzi, A., S. Campanaro, M. D'angelo, F. Simonato, N. Vitulo, F. Lauro, A. Cestaro, G. Malacrida, B. Simionati, N. Cannata, C. Romualdi, D. H. Bartlett, and G. Valle (2005). Life at depth : *Photobacterium profundum* genome sequence and expression analysis. *Science* 307(5714), 1459–1461.
- Vicente, M., A. I. Rico, R. Martínez-Arteaga, and J. Mingorance (2006). Septum enlightenment : assembly of bacterial division proteins. *Journal of Bacteriology* 188(1), 19–27.
- Villaseñor, T., S. Brom, A. Dávalos, L. Lozano, D. Romero, and A. García-de Los Santos (2011). Housekeeping genes essential for pantothenate biosynthesis are plasmidencoded in *Rhizobium etli* and *Rhizobium leguminosarum*. *BMC microbiology* 11(1), 66.
- Vogelmann, J. and M. Ammelburg (2011). Conjugal plasmid transfer in *Streptomyces* resembles bacterial chromosome segregation by FtsK/SpoIIIE. *The EMBO Journal* 30(11), 2246–2254.

- Volff, J.-N. and J. Altenbuchner (2000). A new beginning with new ends : linearisation of circular chromosomes during bacterial evolution. *FEMS Microbiology Letters* 186(2), 143–150.
- Vuilleumier, S., L. Chistoserdova, M.-C. Lee, F. Bringel, A. Lajus, Y. Zhou, B. Gourion, V. Barbe, J. Chang, S. Cruveiller, C. Dossat, W. Gillett, C. Gruffaz, E. Haugen, E. Hourcade, R. Levy, S. Mangenot, E. Muller, T. Nadalig, M. Pagni, C. Penny, R. Peyraud, D. G. Robinson, D. Roche, Z. Rouy, C. Saenampechek, G. Salvignol, D. Vallenet, Z. Wu, C. J. Marx, J. A. Vorholt, M. V. Olson, R. Kaul, J. Weissenbach, C. Médigue, M. E. Lidstrom (2009). *Methylobacterium* genome sequences : a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. *PLoS One* 4(5), e5584.
- Wachi, M., M. Doi, Y. Okada, and M. Matsuhashi (1989). New mre genes mreC and mreD, responsible for formation of the rod shape of Escherichia coli cells. Journal of Bacteriology 171(12), 6511–6516.
- Wang, H., K. Sivonen, L. Rouhiainen, D. P. Fewer, C. Lyra, A. Rantala-Ylinen, J. Vestola, J. Jokela, K. Rantasärkkä, Z. Li, and B. Liu (2012). Genome-derived insights into the biology of the hepatotoxic bloom-forming cyanobacterium *Anabaena* sp. strain 90. *BMC Genomics* 13, 613.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal* of the American Statistical Association 58(301), 236–244.
- Ward Jr, J. H. and M. E. Hook (1963). Application of an hierarchial grouping procedure to a problem of grouping profiles. *Educational and Psychological Measurement* 23(1), 69–81.
- Wattam, A. R., D. Abraham, O. Dalay, T. L. Disz, T. Driscoll, J. L. Gabbard, J. J. Gillespie, R. Gough, D. Hix, R. Kenyon, D. Machil, C. Mao, E. K. Nordberg, R. Olson, R. Overbeek, G. D. Pusch, M. Shukla, J. Schulman, R. L. Stevens, D. E. Sullivan, V. Vonstein, A. Warren, R. Will, M. J.C. Wilson, H. S. Yoo, C. Zhang, Y. Zhang, and B. W. Sobral (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research* 42(Database issue), D581–D591.
- Wattam, A. R., K. P. Williams, E. E. Snyder, N. F. Almeida, M. Shukla, A. W. Dickerman, O. R. Crasta, R. Kenyon, J. Lu, J. M. Shallom, H. Yoo, T. A. Ficht, R. M. Tsolis, C. Munk, R. Tapia, C. S. Han, J. C. Detter, D. Bruce, T. S. Brettin, B. W. Sobral, S. M. Boyle, and J. A. C. Setubal (2009). Analysis of ten *Brucella* genomes reveals evidence for horizontal gene transfer despite a preferred intracellular lifestyle. *Journal of bacteriology 191*(11), 3569–3579.
- Welsh, E. A., M. Liberton, J. Stöckel, T. Loh, T. Elvitigala, C. Wang, A. Wollam, R. S. Fulton, S. W. Clifton, J. M. Jacobs, R. Aurora, B. K. Ghosh, L. A. Sherman, R. D. Smith, R. K. Wilson, and H. B. Pakrasi (2008). The genome of *Cyanothece* 51142, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. *Proceedings of the National Academy of Sciences U.S.A.* 105(39), 15094–15099.

- Wendel, J. and B. P. Buttenfield (2010). Formalizing Guidelines for Building Meaningful Self-Organizing Maps. In GIScience 2010 Short Paper Proceedings, Zurich, Switzerland, September.
- White, O., J. A. Eisen, J. Heidelberg, E. Hickey, J. Peterson, R. Dodson, D. Haft, M. Gwinn, W. Nelson, D. Richardson, K. Moffat, H. Qin, L. Jiang, W. Pamphile, M. Crosby, M. Shen, J. Vamathevan, P. Lam, L. McDonald, T. Utterback, C. Zalewski, K. Makarova, L. Aravind, M. Daly, K. Minton, R. Fleischmann, K. Ketchum, K. Nelson, S. Salzberg, H. Smith, J. Venter, and C. M. C. Fraser (1999). Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* 286(5444), 1571– 1577.
- Wiedenbeck, J. and F. M. Cohan (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiology Reviews* 35(5), 957–976.
- Wisniewski-Dyé, F., K. Borziak, G. Khalsa-Moyers, G. Alexandre, L. O. Sukharnikov, K. Wuichet, G. B. Hurst, W. H. McDonald, J. S. Robertson, V. Barbe, A. Calteau, Z. Rouy, S. Mangenot, C. Prigent-Combaret, P. Normand, M. Boyer, P. Siguier, Y. Dessaux, C. Elmerich, G. Condemine, G. Krishnen, I. Kennedy, A. H. Paterson, V. González, P. Mavingui, and I. B. Zhulin (2011). Azospirillum genomes reveal transition of bacteria from aquatic to terrestrial environments. PLoS Genetics 7(12), e1002430.
- Wistrand, M. and E. L. Sonnhammer (2005). Improved profile HMM performance by assessment of critical algorithmic features in SAM and HMMER. *BMC Bioinformatics* 6(1), 99.
- Witten, I. H., E. Frank, and M. A. Hall (2011). Data Mining : Practical machine learning tools and techniques, Third Edition. Morgan Kaufmann.
- Wolf, Y. I., I. B. Rogozin, A. S. Kondrashov, and E. V. Koonin (2001). Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. *Genome Research* 11(3), 356–372.
- Wollman, E.-L., F. Jacob, and W. Hayes (1956). Conjugation and genetic recombination in Escherichia coli K-12. Cold Spring Harbor symposia on quantitative biology 21, 141–162.
- Wood, D. W., J. C. Setubal, R. Kaul, D. E. Monks, J. P. Kitajima, V. K. Okura, Y. Zhou,
 L. Chen, G. E. Wood, N. F. Almeida Jr., L. Woo, Y. Chen, I. T. Paulsen, J. A. Eisen,
 P. D. Karp, D. Bovee Sr., P. Chapman, J. Clendenning, G. Deatherage, W. Gillet,
 C. Grant, T. Kutyavin, R. Levy, M. -J. Li, E. McClelland, A. Palmieri, C. Raymond,
 G. Rouse, C. Saenphimmachak, Z. Wu, P. Romero, D. Gordon, S. Zhang, H. Yoo,
 Y. Tao, P. Biddle, M. Jung, W. Krespan, M. Perry, B. Gordon-Kamm, L. Liao, S. Kim,
 C. Hendrick, Z. -Y. Zhao, M. Dolan, F. Chumley, S. V. Tingey, J. -F. Tomb, M. P.
 Gordon, M. V. Olson, and E. W. Nester (2001). The genome of the natural genetic
 engineer Agrobacterium tumefaciens C58. Science 294 (5550), 2317–2323.

- Worning, P., L. J. Jensen, P. F. Hallin, H.-H. Stærfeldt, and D. W. Ussery (2006). Origin of replication in circular prokaryotic chromosomes. *Environmental Microbiology* 8(2), 353–361.
- Wozniak, R. a. F. and M. K. Waldor (2010). Integrative and conjugative elements : mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature Reviews Microbiology* 8(8), 552–563.
- Wray, G. A. and E. Abouheif (1998). When is homology not homology? Current Opinion in Genetics & Development 8(6), 675–680.
- Wu, T.-F., C.-J. Lin, and R. C. Weng (2004). Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research* 5(975-1005), 4.
- Xu, Y. (2010). Synechococcus sp. PCC 7002. Ph.D. thesis, The Pennsylvania State University.
- Yamaichi, Y., M. A. Fogel, and M. K. Waldor (2007). par genes and the pathology of chromosome loss in Vibrio cholerae. Proceedings of the National Academy of Sciences U.S.A. 104(2), 630–635.
- Yamaichi, Y., M. A. Gerding, B. M. Davis, and M. K. Waldor (2011). Regulatory cross-talk links *Vibrio cholerae* chromosome II replication and segregation. *PLoS Genetics* 7(7), e1002189.
- Yamaichi, Y., T. Iida, K.-S. Park, K. Yamamoto, and T. Honda (1999). Physical and genetic map of the genome of *Vibrio parahaemolyticus* : presence of two chromosomes in Vibrio species. *Molecular Microbiology* 31(5), 1513–1521.
- Yamanaka, K., T. Ogura, H. Niki, and S. Hiraga (1996). Identification of two new genes, mukE and mukF, involved in chromosome partitioning in Escherichia coli. Molecular and General Genetics 250(3), 241–251.
- Yeoman, C. J., W. J. Kelly, J. Rakonjac, S. C. Leahy, E. Altermann, and G. T. Attwood (2011). The large episomes of *Butyrivibrio proteoclasticus* B316T have arisen through intragenomic gene shuttling from the chromosome to smaller *Butyrivibrio*specific plasmids. *Plasmid* 66(2), 67–78.
- Yoshikawa, H. and N. Ogasawara (1991). Structure and function of DnaA and the DnaA-box in eubacteria : evolutionary relationships of bacterial replication origins. *Molecular Microbiology* 5(11), 2589–2597.
- Yuan, J., Y. Yamaichi, and M. K. Waldor (2011). The three Vibrio cholerae chromosome II-encoded ParE toxins degrade chromosome I following loss of chromosome II. Journal of Bacteriology 193(3), 611–619.
- Zakrzewska-Czerwińska, J., D. Jakimowicz, A. Zawilak-Pawlik, and W. Messer (2007). Regulation of the initiation of chromosomal replication in bacteria. *FEMS Microbiology Reviews* 31(4), 378–387.

- Zheng, J., D. Peng, L. Ruan, and M. Sun (2013). Evolution and dynamics of megaplasmids with genome sizes larger than 100 kb in the *Bacillus cereus* group. *BMC Evolutionary Biology* 13(1), 262.
- Zuerner, R. L., J.-L. Herrmann, and I. Saint Girons (1993). Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity. *Journal of Bacteriology* 175(17), 5445–5451.
Annexe A

Diversité des réplicons secondaires accessoires

TABLE A.1: Diversité des réplicons secondaires essentiels.

ND: Non-Disponible. RefSeq: numéros d'accession RefSeq (NCBI) des réplicons essentiels. taille, ng et forme: taille (en paires de bases), nombres de gènes et configuration (circulaire, linéaire ou ND) des réplicons, respectivement.

	Réplicon essentiel principal				Réplicon essentiel secondaire			ndaire
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
Acidobacteria								
Acidobacteria								
Candidatus Chloracid obacterium thermophilum ${\cal B}$	NC_016024	2683362	2227	circulaire	NC_016025	1012010	826	circulaire
Actinobacteria								
Actinobacteridae								
$No cardiopsis\ dassonvillei\$ subsp. dassonvillei\ DSM 43111	NC_014210	5767958	4798	circulaire	NC_014211	775354	699	circulaire

	Réplic	on essenti	el prin	icipal	Réplicon essentiel seconda			
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
BACTEROIDETES								
Bacteroidia								
Prevotella dentalis DSM 3688	NC_019960	1890695	1412	circulaire	NC_019968	1450390	1121	circulaire
Prevotella intermedia 17	NC_017860	579647	470	circulaire	NC_017861	2119790	1796	circulaire
Prevotella melaninogenica ATCC 25845	NC_014370	1796408	1338	circulaire	NC_014371	1371874	958	circulaire
Chloroflexi								
Chloroflexi non classée								
Thermobaculum terrenum ATCC BAA-798	NC_013525	2026947	1859	circulaire	NC_013526	1074634	973	circulaire
Thermomicrobia								
Sphaerobacter thermophilus DSM 20745	NC_013523	2741033	2439	circulaire	NC_013524	1252731	1046	circulaire
Cyanobacteria								
Hormogoneae								
Anabaena sp. 90	$NC_{-}019427$	4329264	3648	circulaire	NC_019439	819965	706	circulaire
Chroobacteria								
Cyanothece sp. ATCC 51142	NC_010546	4934271	4761	circulaire	NC_010547	429701	449	linéaire
Deinococcus-Thermus								
Deinococci								
Deinococcus radiodurans R1	NC_001263	2648638	2629	circulaire	NC_001264	412348	368	circulaire
FIRMICUTES								
Clostridia								
Butyrivibrio proteoclasticus B316	NC_014387	3554804	2939	circulaire	NC_014388	302358	250	circulaire
Clostridium difficile BI1	NC_017179	4118573	3591	circulaire	NC_017177	300869	279	circulaire

	Réplicon essentiel principal			Réplicon essentiel secondaire				
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
Fusobacteria								
Fusobacteriia								
Ilyobacter polytropus DSM 2926	NC_014632	2046464	1889	circulaire	NC_014633	961624	882	circulaire
Proteobacteria								
Alphaproteobacteria								
Agrobacterium fabrum C58	NC_003062	2841580	2765	circulaire	NC_003063	2075577	1851	linéaire
Agrobacterium radiobacter K84	NC_011985	4005130	3744	circulaire	NC_011983	2650913	2363	circulaire
Agrobacterium sp. H13-3	NC_015183	2823930	2752	circulaire	NC_015508	2148289	1942	linéaire
Agrobacterium vitis S4	NC_011989	3726375	3234	circulaire	NC_011988	1283187	1054	circulaire
Asticcacaulis excentricus CB 48	NC_014816	2588221	2330	circulaire	NC_014817	1315949	1121	circulaire
Brucella abortus A13334	NC_016795	2123773	2185	circulaire	NC_016777	1162259	1153	circulaire
Brucella abortus S19	NC_010742	2122487	1967	circulaire	NC_010740	1161449	1033	irculaire
Brucella abortus bv. 1 str. 9-941	NC_006932	2124241	2029	circulaire	NC_006933	1162204	1055	circulaire
Brucella canis ATCC 23365	NC_010103	2105969	2100	circulaire	NC_010104	1206800	1149	circulaire
Brucella canis HSK A52141	NC_016778	2107023	2129	circulaire	NC_016796	1170489	1151	circulaire
Brucella melitensis ATCC 23457	NC_012441	2125701	2063	circulaire	NC_012442	1185518	1072	circulaire
Brucella melitensis M28	NC_017244	2126133	2173	circulaire	NC_017245	1185615	1190	circulaire
Brucella melitensis M5-90	NC_017246	2126451	2170	circulaire	NC_017247	1185778	1187	circulaire
Brucella melitensis NI	NC_017248	2117717	2081	circulaire	NC_017283	1176758	1148	circulaire
Brucella melitensis bv. Abortus 2308	NC_007618	2121359	2000	circulaire	NC_007624	1156948	1034	circulaire
Brucella melitensis bv. 1 str. 16M	NC_003317	2117144	2059	circulaire	NC_003318	1177787	1139	circulaire
Brucella microti CCM 4915	NC_013119	2117050	2115	circulaire	NC_013118	1220319	1167	circulaire
Brucella ovis ATCC 25840	NC_009505	2111370	1928	circulaire	NC_009504	1164220	961	circulaire

	Réplico	Réplicon essentiel principal			Réplicon essentiel secondaire			
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
Brucella pinnipedialis B2/94	NC_015857	2138342	2136	circulaire	NC_015858	1260926	1188	circulaire
Brucella suis 1330	NC_004310	2107794	2122	circulaire	NC_004311	1207381	1149	circulaire
Brucella suis 1330	NC_017251	2107783	2122	circulaire	NC_017250	1207380	1144	circulaire
Brucella suis ATCC 23445	NC_010169	1923763	1920	circulaire	NC_010167	1400844	1320	circulaire
Brucella suis VBI22	NC_016797	2108637	2124	circulaire	NC_016775	1207451	1146	circulaire
Ochrobactrum anthropi ATCC 49188	NC_009667	2887297	2731	circulaire	NC_009668	1895911	1693	circulaire
Paracoccus denitrificans PD1222	NC_008686	2852282	2799	circulaire	NC_008687	1730097	1662	circulaire
Rhizobium sp. IRBG74	NC_022535	2844565	2882	circulaire	NC_022545	2035452	1921	linéaire
Rhodobacter sphaeroides 2.4.1	NC_007493	3188524	3021	circulaire	NC_007494	943018	835	circulaire
Rhodobacter sphaeroides ATCC 17025	NC_009428	3217726	3111	circulaire	NC_009429	877879	802	circulaire
Rhodobacter sphaeroides ATCC 17029	NC_009049	3147721	2972	circulaire	NC_009050	1219053	1052	circulaire
Rhodobacter sphaeroides KD131	NC_011963	3152792	3100	circulaire	NC_011958	1297647	1224	circulaire
Rhodobacter sphaeroides WS8N	NZ_CM001161	ND	ND	ND	NZ_CM001162	ND	ND	ND
Sinorhizobium meliloti AK83	NC_015590	3820344	3626	circulaire	NC_015591	1312480	1136	circulaire
					NC_015596	1680879	1491	circulaire
Sphingobium chlorophenolicum L-1	NC_015593	3080818	2846	circulaire	NC_015594	1368670	1104	circulaire
Sphingobium japonicum UT26S	NC_014006	3514822	3529	circulaire	NC_014013	681892	589	circulaire
Betaproteobacteria								
Burkholderia ambifaria AMMD	NC_008390	3556545	3206	circulaire	NC_008391	2646969	2346	circulaire
					NC_008392	1281472	1013	circulaire
Burkholderia ambifaria MC40-6	NC_010551	3443583	3074	circulaire	NC_010552	2769414	2382	circulaire
					NC_010557	1127947	967	circulaire
Burkholderia cenocepacia AU 1054	NC_008060	3294563	2965	circulaire	NC_008061	2788459	2472	circulaire
					NC_008062	1196094	1040	circulaire
Burkholderia cenocepacia HI2424	NC_008542	3483902	3159	circulaire	NC_008543	2998664	2686	circulaire

	Réplicon essentiel principal			Réplicon essentiel secondaire				
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
					NC_008544	1055417	918	circulaire
Burkholderia cenocepacia J2315	NC_011000	3870082	3464	circulaire	NC_011001	3217062	2807	circulaire
					NC_011002	875977	752	circulaire
Burkholderia cenocepacia MC0-3	NC_010508	3532883	3160	circulaire	NC_010512	1224595	1053	circulaire
					$NC_{-}010515$	3213911	2795	circulaire
Burkholderia cepacia GG4	NC_018513	3463655	3120	circulaire	NC_018514	3003666	2705	circulaire
Burkholderia gladioli BSR3	NC_015381	4413616	3779	circulaire	NC_015376	3700833	2926	circulaire
Burkholderia glumae BGR1	NC_012724	3906507	3287	circulaire	NC_012721	2827333	2079	circulaire
Burkholderia lata	NC_007510	3694126	3334	circulaire	NC_007509	1395069	1209	circulaire
					NC_007511	3587082	3173	circulaire
Burkholderia mallei ATCC 23344	NC_006348	3510148	2995	circulaire	NC_006349	2325379	2027	circulaire
Burkholderia mallei NCTC 10229	NC_008836	3458208	3332	circulaire	NC_008835	2284095	2177	circulaire
Burkholderia mallei NCTC 10247	NC_009080	3495687	3274	circulaire	NC_009079	2352693	2141	circulaire
Burkholderia mallei SAVP1	NC_008785	3497479	3454	circulaire	NC_008784	1734922	1730	circulaire
Burkholderia multivorans ATCC 17616	NC_010804	3448421	3084	circulaire	NC_010801	919805	767	circulaire
					NC_010805	2473162	2129	circulaire
Burkholderia multivorans ATCC 17616	NC_010084	3448466	3146	circulaire	NC_010086	2472928	2151	circulaire
					NC_010087	919806	823	circulaire
Burkholderia phenoliruptrix BR3459a	NC_018695	4152217	3491	circulaire	NC_018672	2713495	2289	circulaire
Burkholderia phymatum STM815	NC_010622	3479187	3072	circulaire	NC_010623	2697374	2349	circulaire
Burkholderia phytofirmans PsJN	NC_010681	4467537	3922	circulaire	NC_010676	3625999	3152	circulaire
Burkholderia pseudomallei 1026b	NC_017831	4092668	3608	circulaire	NC_017832	3138747	2462	circulaire
Burkholderia pseudomallei 1106a	NC_009076	3988455	4015	circulaire	NC_009078	3100794	3159	circulaire
Burkholderia pseudomallei 1106b	NZ_CM000774	ND	ND	ND	NZ_CM000775	ND	ND	ND
Burkholderia pseudomallei 1710a	NZ_CM000832	ND	ND	ND	NZ_CM000833	ND	ND	ND

	Réplicon essentiel principal			Réplicon	Réplicon essentiel secondaire			
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
Burkholderia pseudomallei 1710b	$NC_{-}007434$	4126292	3733	circulaire	NC_007435	3181762	2611	circulaire
Burkholderia pseudomallei 668	NC_009074	3912947	3939	circulaire	NC_009075	3127456	3177	circulaire
Burkholderia pseudomallei BPC006	NC_018527	4001777	4050	circulaire	NC_018529	3153284	3108	circulaire
Burkholderia pseudomallei K96243	NC_006350	4074542	3399	circulaire	NC_006351	3173005	2329	circulaire
Burkholderia pseudomallei MSHR305	NC_021884	4054155	3461	circulaire	NC_021877	3373917	2644	circulaire
Burkholderia sp. CCGE1001	$NC_{-}015136$	4063449	3545	circulaire	NC_015137	2770302	2420	circulaire
Burkholderia sp. CCGE1002	NC_014117	3518940	3116	circulaire	NC_014118	2593966	2258	circulaire
					NC_014119	1282816	1109	circulaire
Burkholderia sp. CCGE1003	NC_014539	4077097	3463	circulaire	NC_014540	2966498	2525	circulaire
Burkholderia sp. KJ006	NC_017920	3145156	2917	circulaire	NC_017921	2356985	2132	circulaire
					NC_017922	1082410	930	circulaire
Burkholderia sp. RPE64	$NC_{-}021287$	3013410	2907	circulaire	NC_021288	900830	853	circulaire
					NC_021294	1465356	1422	circulaire
Burkholderia sp. YI23	NC_016589	3131280	2769	circulaire	NC_016590	1569570	1364	circulaire
					NC_016625	1773019	1539	circulaire
Burkholderia thailandensis E264	$NC_{-}007651$	3809201	3276	circulaire	NC_007650	2914771	2356	circulaire
Burkholderia thailandensis E264	NZ_CM000438	ND	ND	ND	NZ_CM000439	ND	ND	ND
Burkholderia thailandensis MSMB121	NC_021173	3967794	3527	circulaire	NC_021174	2763585	2231	circulaire
Burkholderia vietnamiensis G4	NC_009256	3652814	3274	circulaire	NC_009254	1241007	1114	circulaire
					NC_009255	2411759	2096	circulaire
Burkholderia xenovorans LB400	$NC_{-}007951$	4895836	4430	circulaire	NC_007952	3363523	2960	circulaire
					NC_007953	1471779	1312	circulaire
Cupriavidus necator N-1	NC_015726	3872936	3622	circulaire	NC_015723	2684606	2443	circulaire
Cupriavidus taiwanensis LMG 19424	NC_010528	3416911	3135	circulaire	NC_010530	2502411	2242	circulaire
Ralstonia eutropha H16	NC_008313	4052032	3651	circulaire	NC_008314	2912490	2555	circulaire

	Réplicon essentiel principal			Réplicon essentiel secondaire				
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
Ralstonia eutropha JMP134	NC_007347	3806533	3439	circulaire	NC_007348	2726152	2407	circulaire
Ralstonia pickettii 12D	NC_012856	3647724	3411	circulaire	NC_012857	1323321	1180	circulaire
Ralstonia pickettii 12J	NC_010682	3942557	3709	circulaire	NC_010678	1302238	1153	circulaire
Variovorax paradoxus B4	NC_022247	5795261	5434	circulaire	NC_022234	1353255	1319	circulaire
Variovorax paradoxus S110	NC_012791	5626353	5241	circulaire	NC_012792	1128644	1038	circulaire
Gammaproteobacteria								
Aliivibrio salmonicida LFI1238	NC_011312	3325165	2820	circulaire	NC_011313	1206461	984	circulaire
Listonella anguillarum M3	NC_022223	3063587	2719	circulaire	NC_022224	988134	900	circulaire
Photobacterium profundum SS9	NC_006370	4085304	3416	circulaire	NC_006371	2237943	2006	circulaire
Pseudoalteromonas haloplanktis TAC125	NC_007481	3214944	2939	circulaire	NC_007482	635328	545	circulaire
Pseudoalteromonas sp. SM9913	NC_014803	3332787	3078	circulaire	NC_014800	704884	634	circulaire
Vibrio alginolyticus NBRC $15630 = \text{ATCC } 17749$	NC_022349	3334467	3054	circulaire	NC_022359	1812170	1644	circulaire
Vibrio anguillarum 775	NC_015633	3063912	2812	circulaire	NC_015637	988135	920	circulaire
Vibrio campbellii ATCC BAA-1116	NC_009783	3765351	3548	circulaire	NC_009784	2204018	2373	circulaire
Vibrio campbellii ATCC BAA-1116	NC_022269	3746887	3327	circulaire	NC_022270	2195939	2001	circulaire
Vibrio cholerae IEC224	NC_016944	3007450	2635	circulaire	NC_016945	1072136	1029	circulaire
Vibrio cholerae LMA3984-4	NC_017270	2791729	2328	circulaire	NC_017269	946986	825	circulaire
Vibrio cholerae M66-2	NC_012578	2892523	2650	circulaire	NC_012580	1046382	964	circulaire
Vibrio cholerae MJ-1236	NC_012668	3149584	2768	circulaire	NC_012667	1086784	1004	circulaire
Vibrio cholerae O1 biovar El Tor str. N16961	NC_002505	2961149	2534	circulaire	NC_002506	1072315	970	circulaire
Vibrio cholerae O1 str. 2010EL-1786	NC_016445	3031375	2805	circulaire	NC_016446	1046365	1040	circulaire
Vibrio cholerae O1 str. Inaba G4222	NZ_CM001785	ND	ND	ND	NZ_CM001786	ND	ND	ND
Vibrio cholerae O395	NC_009457	3024069	2742	circulaire	NC_009456	1108250	1133	circulaire
Vibrio cholerae O395	NC_012582	3024078	2817	circulaire	NC_012583	1111222	1117	circulaire

	Réplicon essentiel principal			Réplicon essentiel secondaire				
Espèce bactérienne	RefSeq	taille	ng	forme	RefSeq	taille	ng	forme
Vibrio fischeri ES114	NC_006840	2897536	2584	circulaire	NC_006841	1330333	1174	circulaire
Vibrio fischeri MJ11	NC_011184	2905029	2590	circulaire	NC_011186	1418848	1254	circulaire
Vibrio fischeri SR5	NZ_CM001400	ND	ND	ND	NZ_CM001401	ND	ND	ND
Vibrio furnissii NCTC 11218	NC_016602	3294546	3006	circulaire	NC_016628	1621862	1449	circulaire
Vibrio nigripulchritudo SnF1	NC_022528	4109740	3570	circulaire	NC_022543	2212415	2004	circulaire
Vibrio parahaemolyticus BB22OP	NC_019955	3297305	2847	circulaire	NC_019971	1806219	1600	circulaire
Vibrio parahaemolyticus RIMD 2210633	NC_004603	3288558	3079	circulaire	NC_004605	1877212	1752	circulaire
Vibrio sp. EJY3	NC_016613	3478307	3075	circulaire	NC_016614	1974339	1711	circulaire
Vibrio sp. Ex25	NC_013456	3259580	2915	circulaire	NC_013457	1829445	1603	circulaire
Vibrio splendidus LGP32	NC_011753	3299303	2947	circulaire	NC_011744	1675515	1485	circulaire
Vibrio vulnificus CMCP6	NC_004459	3281866	2896	circulaire	NC_004460	1844830	1537	circulaire
Vibrio vulnificus MO6-24/O	NC_014965	3194232	2980	circulaire	NC_014966	1813536	1582	circulaire
Vibrio vulnificus VVyb1(BT3)	NZ_CM001799	ND	ND	ND	NZ_CM001800	ND	ND	ND
Vibrio vulnificus YJ016	NC_005139	3354505	3259	circulaire	NC_005140	1857073	1695	circulaire
Spirochaetae								
Spirochaetes								
Leptospira biflexa serovar Patoc strain 'Patoc 1 (Ames)'	NC_010842	3603977	3277	circulaire	$NC_{-}010845$	277995	266	circulaire
Leptospira biflexa serovar Patoc strain 'Patoc 1 (Paris)'	$NC_{-}010602$	3599677	3390	circulaire	NC_010843	277655	276	circulaire
Leptospira borgpetersenii serovar Hardjo-bovis str. JB197	$NC_{-}008510$	3576473	2645	circulaire	NC_008511	299762	235	circulaire
$Leptospira\ borgpetersenii\ serovar$ Hardjo-bovis str. L 550	$NC_{-}008508$	3614446	2703	circulaire	$NC_{-}008509$	317336	242	circulaire
Leptospira interrogans serovar Copenhageni str. Fiocruz L1-130	NC_005823	4277185	3399	circulaire	NC_005824	350181	268	circulaire
Leptospira interrogans serovar Lai str. 56601	NC_004342	4338762	3409	circulaire	NC_004343	359372	293	circulaire
Leptospira interrogans serovar Lai str. IPAV	NC_017551	4349158	3418	circulaire	NC_017552	359372	293	circulaire

Annexe B

Groupes d'orthologie KEGG sélectionnés

 TABLE B.1: Groupes d'orthologie sélectionnés à partir de la hiérarchie KEGG-BRITE.

Colonne 1 : identifiant KEGG de la famille d'orthologie, Colonne 2 : gène(s) représentatif(s) (cf. Chapitre 1), Colonne 3 : description succincte de la fonction.

K03746	hns	DNA-binding protein H-NS						
K11685	stpA	DNA-binding protein StpA						
HU (hea	t unstable protein)						
K05787	hupA	DNA-binding protein HU-alpha						
K03530	hupB	DNA-binding protein HU-beta						
IHF (integration host factor)								
K04764	<i>ihfA</i> , <i>himA</i> integration host factor subunit alpha							
K05788	ihfB, himD	integration host factor subunit beta						
Other nu	ner nucleoid associated proteins							
K05516	cbpA	curved DNA-binding protein						
K05804	rob	right origin-binding protein						
K02313	dnaA	chromosomal replication initiator protein						
K12961	diaA	DnaA initiator-associating protein						
K04047	dps	starvation-inducible DNA-binding protein						
K03557	fis	Fis family transcriptional regulator, factor for inversion						
		stimulation protein						
K03666	hfq	host factor-I protein						
K05596	iciA	LysR family transcriptional regulator, chromosome initiation inhibitor						
K03719	lrp	Lrp/AsnC family transcriptional regulator, leucine- responsive regulatory protein						

HNS (histone-like nucleoid structuring protein)

MUKDE	complex	
K03632	mukB	chromosome partition protein MukB
K03804	mukE	chromosome partition protein MukE
K03633	mukF	chromosome partition protein MukF
Condens	in-like complex	
$\mathbf{K03529}$	smc	chromosome segregation protein
K05896	scpA	segregation and condensation protein A
$\mathbf{K06024}$	scpB	segregation and condensation protein B
Divisome	e proteins	
K03590	ftsA	cell division protein FtsA
K05589	ftsB	cell division protein FtsB
K13052	divIC, divA	cell division protein DivIC
K09812	ftsE	cell division transport system ATP-binding protein
K03587	ftsI	cell division protein FtsI (penicillin-binding protein 3) [EC :2.4.1.129]
K03466	ftsK, spoIIIE	DNA segregation ATPase FtsK/SpoIIIE, S-DNA-T family
K03586	ftsL	cell division protein FtsL
K03591	ftsN	cell division protein FtsN
K03589	ftsQ	cell division protein FtsQ
K03588	ftsW, spoVE	cell division protein FtsW
K09811	ftsX	cell division transport system permease protein
K03531	ftsZ	cell division protein FtsZ
K03528	zipA	cell division protein ZipA
K09888	zapA	cell division protein ZapA
K01448	amiA, amiB, amiC	N-acetylmuramoyl-L-alanine amidase [EC :3.5.1.28]
K03585	acrA	membrane fusion protein
Inhibitor	s of FtsZ assembl	У
K06286	ezrA	septation ring formation regulator
K03610	minC	septum site-determining protein MinC
K03609	minD	septum site-determining protein MinD
K03608	minE	cell division topological specificity factor
$\mathbf{K04074}$	divIVA	cell division initiation protein
K13053	sulA	cell division inhibitor SulA
K09772	sepF	cell division inhibitor SepF
K05501	slmA, ttk	TetR/AcrR family transcriptional regulator
Other ch	romosome partiti	oning proteins
K03645	seqA	negative modulator of initiation of replication

Chromosome partitioning proteins MukBEF complex

K03645	seqA	negative modulator of initiation of replication
K03569	mreB	rod shape-determining protein MreB and related proteins
K03570	mreC	rod shape-determining protein MreC
K03571	mreD	rod shape-determining protein MreD
K05837	rodA, mrdB	rod shape determining protein RodA

K03496	parA, soj	chromosome partitioning protein						
K03497	parB, spo0J	chromosome partitioning protein, ParB family						
K02621	parC	topoisomerase IV subunit A [EC :5.99.1]						
K02622	parE	topoisomerase IV subunit B [EC :5.99.1]						
K03495	gidA, mnmG,	tRNA uridine 5-carboxymethylaminomethyl modification						
	MTO1	enzyme						
K03501	$gidB, \ rsmG$	16S rRNA (guanine527-N7)-methyltransferase						
		[EC : 2.1.1.170]						
K04094	trmFO, gid	methylenetetrahydrofolate-tRNA-(uracil-5-)-						
		methyltransferase $[EC : 2.1.1.74]$						
K03733	xerC	integrase/recombinase XerC						
K04763	xerD	integrase/recombinase XerD						
K11686	racA	chromosome-anchoring protein RacA						
K03593	mrp	ATP-binding protein involved in chromosome partitioning						
K04095	fic	cell filamentation protein						
Initiation	n factors (bacteria	1)						
K05787	hupA	DNA-binding protein HU-alpha						
K03530	hupB	DNA-binding protein HU-beta						
K04764	ihfA, himA	integration host factor subunit alpha						
K05788	ihfB, himD	integration host factor subunit beta						
K02313	dnaA	chromosomal replication initiator protein						
K02314	dnaB	replicative DNA helicase [EC :3.6.4.12]						
K02315	dnaC	DNA replication protein DnaC						
K02316	dnaG	DNA primase [EC :2.7.7]						
K03346	$dnaB2, \ dnaB$	replication initiation and membrane attachment protein						
K11144	dnaI	primosomal protein DnaI						
K03111	ssb	single-strand DNA-binding protein						
Terminus	s site-binding prot	tein						
K10748	tus, tau	DNA replication terminus site-binding protein						

DNA methylation enzyme

K06223	dam	DNA adenine methylase [EC :2.1.1.72]
Preventi	on of re-replicatio	n factors
K03645	seqA	negative modulator of initiation of replication
K10763	hda	DnaA-homolog protein

Annexe C

Familles protéiques ACLAME sélectionnées

Présentation des familles de protéines sélectionnées dans ACLAME (v0.4) pour les analyses du Chapitre 3.

Famille	Nombre	Description	Annotation
32	185	RepB, pi, initiator protein, RepE, RepA	RepA, E, B
76	89	Rep, RepB, Rep of rolling circle initiator, RepA, RepU, OrfB, Rep2	RepB
107	61	$\operatorname{RepC}, \operatorname{RepCa1}, \operatorname{RepCa2}, \operatorname{RepCd}$	RepC
114	59	Helicase, UrvD rep helicase, helicase superfamily1, Yga2F, helicase II	Helicase
118	59	CdsE, CdsJ	Cds (Borrelia)
133	51	$\operatorname{RepA}, \operatorname{W0005}, \operatorname{RepA1/A2}$	RepA
171	44	$\operatorname{Rep} A,$ $\operatorname{Rep} B,$ putative theta replicative protein	RepA
207	39	replicative DNA helicase, DnaB, pGP1	DnaB
208	39	$\operatorname{RepA}, \operatorname{W0013}, \operatorname{W0041}, \operatorname{RepFIB}$	RepA
224	37	long form TrfA, TrfA1, TrfA2, S-TrfA	TrfA
237	36	RepA, putative RepA, truncated RepA	RepA
244	34	$\operatorname{RepA}, \operatorname{RepB}, \operatorname{CopB}, \operatorname{repA1/A2}, \operatorname{w0004}$	RepA, B CopB
294	29	plasmid copy number control, Rop protein, RNAI modulator, RNA modulator	Rop
297	29	primase activity/DNA initiation, LtrC/LtrC-like hypothetical protein, PcfD	primase, LtrC
330	26	DNA repair/ DNA helicase, type III restriction enzyme, res subunit, DEAD/DEAH box helicase	DNA helicase
377	23	replicase, replication initiation, RepC, RepJ, RepE, RepL	$\operatorname{RepC}, J, E$
383	23	RepA , $\operatorname{Rb100}$	RepA
404	22	RepA, RepB, RepW	RepA
412	22	Rep, RepA	Rep

TABLE C.1: Familles ACLAME liées à la réplication.

Famille	Nombre	Description	Annotation
423	21	$truncated \ RCR \ replication, \ RepRC, \ RepB, \ OrfA$	Rep
426	21	cell division control protein 6 homolog	cdc6 archee
440	20	${\rm Rep}14\text{-}4, {\rm rm}\; {\rm protein}, {\rm RepA}\; {\rm hypothetical}\; {\rm protein}$	RepA
451	20	RepA, host type : Corynebacterium	RepA
477	19	Rep, RepS, RepE, host type : $Bacillus$, RepS, RepR	RepR,S,E
612	15	RepL, replication initiation	RepA
775	12	DNA helicase activity, RepA, putative helicase	RepC
854	11	DNA helicase activity, RepC, putative initiator protein	RepC
921	11	RepA	RepA
931	11	DNA replication initiation, putative protein, CdsD	cdsD
1005	10	helicase activity, putative protein, hypothetical helicase	helicase
1055	10	RNA polymerase sigma factor, sigma 70 family, bacteriocin uviA, sigF/V/G, tetR transcriptional activator, host type : $Clostridium$	RNA polymerase
1095	9	DNA repair/ helicase, holliday junction helicase RuvB, DNA pol III gamma and tau subunits, DNA pol delta subunit	RuvB
1099	9	putative theta replicase, RepB, Rep2	RepB
1187	9	DNA replication, RepH, RepI	RepH
1288	8	RepA	RepA
1345	8	DNA primase activity, DNA primase, primase CHC2 family	DNA primase
1398	γ	helicase activity, GcrE, GcrC	helicase
1652	$\tilde{\gamma}$	DNA repair/exonuclease activity, DNA exonuclease protein, SbcCD related protein	DNA repair
1837	5	putative replication protein	Rep
2881		RepC-like, Pif	RepC

TABLE C.2: Familles ACLAME liées à la ségrégation.

Famille	Nombre	Description	Annotation
4	585	plasmid partition protein, ParA, ParA IncC protein, ParA InC1/ IncC2, SopA, virC1	ParA/ParM
14	265	RepB, RepB partitionning, KorB repressor and partitionning, ParB-like domain, YefA, YdeB, ParB, ParB-like	ParB
102	64	DNA binding, partitionning protein, control pro- tein, ParB, VirB, patrition protein B	ParB
128	54	DNA segregation/DNA translocase activity, cell division FtsK/ SpoIIIE, SpoI, TraB	FtsK/SpoIIIE

Famille	Nombre	Description	Annotation
289	29	ParM family, go : translocase, hypothetical pro- tein, rode shape protein, putative ATPase of class HSP70	ParA/ParM
316	27	microfilament motor activity, ParM family, StbA protein, stable inheritance protein, ParA	ParA/ParM
318	27	ATPase/regulation of cell division, ATPase in- volved in chromosome partition, GumC, ExoP related protein, EpsB, MPA1 family	ATPase/tyrK/exoP
427	21	ATPase family, ParR family, ParB, StbB, me- diator of plasmid stability	ParR,ParB
875	11	DNA binding, partitionning protein family ParB/Spo0J, YPMT1.28c	ParB
876	11	DNA binding, partitionning protein family ParB/Spo0J, YPMT1.29c	ParB
983	10	DNA binding, ParB, CopG	ParB
1227	8	DNA plasmid copy number control, CopG	CopG
2158	5	RepC	RepC
2894	4	DNA binding protein	DNA binding

TABLE C.3: Familles ACLAME liées à la résolution de dimères.

Famille	Nombre	Description	Annotation
5	493	serine based recombinase activity, ylb, resolvase, $% \left({{{\bf{x}}_{i}}} \right)$	serine recombinase
		second invertase, TniR, ParA	
10	355	tyrosine-based recombinase, integrase, putative	Xer Tyrosine
		integrase, Xer, recombinase-like SAM	
101	64	plasmid dimer resolution, tyrosine-based recom-	PL dimer resolu-
		binase, yld, SAM—-like protein	tion
170	45	tyrosine-based recombinase, OrfA	Tyrosine rec OrfA
589	15	tyrosine based protein, Fis protein	Fis
688	15	tyrosine based protein, SAM like protein, XerD	XerD

TABLE C.4: Familles ACLAME liés à la maintenance.

Famille	Nombre	Description	Annotation
100	64	Postsegregational killing system vapBC/vag $$	PSK_vapBC/vag
136	50	Postsegregational killing system parDE	PSK_parDE
156	46	Postsegregational killing system epsilon-zeta	$PSK_{epsilon-zeta}$
201	39	Postsegregational killing system higBA	PSK_higBA
212	38	Postsegregational killing system parDE	PSK_parDE
293	29	Postsegregational killing system mazEF	PSK_mazEF
319	26	Postsegregational killing system relBE	PSK_relBE
326	26	Postsegregational killing system mazEF	PSK_mazEF
335	26	Postsegregational killing system HOK/SOK	PSK_HOK/SOK

Famille	Nombre	Description	Annotation
338	25	Postsegregational killing system parDE	PSK_parDE
356	24	Postsegregational killing system parDE	PSK_parDE
366	24	Postsegregational killing system vapBC/vag $$	$\mathrm{PSK}_{-}\mathrm{vapBC}/\mathrm{vag}$
380	23	Postsegregational killing system phD-doc	$\mathrm{PSK}_{-}\mathrm{phD}\text{-}\mathrm{doc}$
428	21	Postsegregational killing system ccd	PSK_ccd
470	19	Postsegregational killing system yacA	PSK_yacA
474	19	Postsegregational killing system relBE	PSK_relBE
515	17	Postsegregational killing system relBE	PSK_relBE
556	16	Postsegregational killing system higBA	PSK_higBA
563	16	Postsegregational killing system ccd	PSK_ccd
588	15	Postsegregational killing system higBA	PSK_higBA
677	14	Postsegregational killing system higBA	PSK_higBA
798	12	Postsegregational killing system mazEF	PSK_mazEF
916	11	Postsegregational killing system relBE	PSK_relBE
1031	10	Postsegregational killing system HOK/SOK	PSK_HOK/SOK
1180	9	Postsegregational killing system vapXD	PSK_vapXD
1308	8	Postsegregational killing system HicAB	PSK_HicAB
1559	$\tilde{\gamma}$	Postsegregational killing system epsilon-zeta	${\rm PSK_epsilon-zeta}$
1927	5	Postsegregational killing system mazEF	PSK_mazEF
3357	3	plasmid maintenance, Postsegregational killing	plasmid mainte-
		system	nance, PSK
4776	2	Postsegregational killing system, parC	PSK, parC
4777	2	Postsegregational killing system parDE, parD	PSK_parDE
16584	1	Postsegregational killing system vapXD	PSK_vapXD

Annexe D

Résultats du clustering de V^R par INFOMAP

Résultats du clustering de V^R par l'algorithme INFOMAP avec une granularité gr de 4 et les données décrites dans le Chapitre 5.

Les réplicons sont groupés par clusters. La valeur de l'indice de stabilité, \triangle^C (eq. 5.14), est donnée pour chaque cluster. Une valeur de \triangle^C inférieure à 0.5 signale un cluster "dissous" et non interprétable, et une valeur supérieure à 0.75 indique un cluster stable [Hennig, 2007, 2008].

Pour chaque réplicon, sont donnés le type (plasmide (**PL**), chromososome (**CHR**) ou RECE avec **RECE*** indiquant les RECE secondaires de *Burkholderia* et *Sinorhizsobium*), le numéro d'accession RefSeq, l'espèce et la classe de la bactérie correspondante et la valeur de l'indice \triangle^e (eq. 5.16) de sa stabilité au sein de son cluster. Pour un réplicon donné, une valeur de \triangle^e égale à "nan" (*not a number*) indique que ce réplicon n'est apparu dans aucune des trente réplication de *bootstrap* de V^R utilisées dans la procédure.

Les RECE sont surlignés en bleu.

Bacilli

Bacilli Bacilli

Bacilli Bacilli

Bacilli

Bacilli

Bacilli

Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli

Bacilli Bacilli

Bacilli

Bacilli Bacilli

Bacilli

Bacilli Bacilli Bacilli

Bacilli Bacilli

Bacilli Bacilli

Bacilli Bacilli

Bacilli Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli Bacilli

Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli

Bacilli Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli Bacilli

Bacilli

Bacilli

Bacilli Bacilli

Bacilli

Bacilli Bacilli

Bacilli Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli

Bacilli Bacilli

Bacilli Bacilli

Bacilli Bacilli

Bacilli

Bacilli Bacilli

Bacilli

Bacilli Bacilli Bacilli

Bacilli

Bacilli Bacilli $1.00 \\ 1.00 \\ 1.00$

1.00

0.90

1.00

nan

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

 $\begin{array}{c} 1.00 \\ 1.00 \end{array}$

1.00

1.00

 $\begin{array}{c} 1.00 \\ 1.00 \end{array}$

1.00

 $1.00 \\ 1.00 \\ 1.00$

 $1.00 \\ 1.00 \\ 1.00$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

nan 1.00

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

nan

 $\begin{array}{c} 1.00 \\ 1.00 \end{array}$

 $1.00 \\ 1.00 \\ 1.00$

1.00

1.00

 $1.00 \\ 1.00$

1.00

 $\begin{array}{c} 1.00\\ 1.00 \end{array}$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

1.00

 $1.00 \\ 1.00$

 $1.00 \\ 1.00 \\ 1.00$

1.00

1.00

 $1.00 \\ 1.00$

1.00

 $1.00 \\ 1.00 \\ 1.00$

1.00

1.00

 $1.00 \\ 1.00$

cluster :	1 stability me	asure : 0.947381			CHR	NC_009328	G. thermodenitrificans
CHR	NC_016024	Cand. Chloracidobacte-	Acidobacteria	0.73	CHR	$NC_{-}015660$	$G.\ thermoglucosidasius$
		rium			CHR	NC_016593	G. thermoleovorans
RECE	NC_016025	Cand. Chloracidobacte-	Acidobacteria	1.00	CHR	NC_017668	H. halophilus
CHD	NC 011567	rium	Desilli	1.00		NC_017669	H. halophilus
CHR	NC 015278	A. jiaviinermus	Bacilli	1.00	CHR	NC 006814	L. acidophilus
CHR	NC_018704	A. xylanus	Bacilli	1.00	CHR	NC_015214	L. acidophilus
CHR	NC_009725	B. amyloliquefaciens	Bacilli	1.00	CHR	NC_014724	L. amylovorus
CHR	NC_014551	B. amyloliquefaciens	Bacilli	1.00	CHR	NC_017470	L. amylovorus
CHR	NC_016784	$B. \ amylolique faciens$	Bacilli	1.00	CHR	$NC_{-}008497$	L. brevis
CHR	NC_017061	B. amyloliquefaciens	Bacilli	1.00	CHR	NC_015428	L. buchneri
CHR	NC_017188	B. amyloliquefaciens	Bacilli	1.00	CHR	NC_018610	L. buchneri
CHR	NC_017190	B. amyloliquefaciens	Bacilli	1.00	CHR	NC_018673	L. carnosum
CHR	NC_017191	B. amyloliquefaciens	Bacilli	nan 1.00	CHR	NC_008526	L. casei
CHB	NC 003997	B. anthracis	Bacilli	1.00	CHB	NC 014334	L casei
CHR	NC 005945	B. anthracis	Bacilli	1.00	CHR	NC_017473	L. casei
CHR	NC_007530	B. anthracis	Bacilli	1.00	CHR	NC_017474	L. casei
CHR	NC_012581	B. anthracis	Bacilli	1.00	CHR	NC_018641	L. casei
CHR	NC_012659	B. anthracis	Bacilli	1.00	CHR	NC_010471	L. citreum
CHR	NC_017729	B. anthracis	Bacilli	1.00	CHR	NC_014106	L. crispatus
CHR	NC_014639	B. atrophaeus	Bacilli	1.00	CHR	NC_008054	L. delbrueckii
CHR	NC_012491	B. brevis	Bacilli	1.00	CHR	NC_008529	L. delbrueckii
CHR	NC_014829	B. cellulosilyticus	Bacilli	1.00	CHR	NC_014727	L. delbrueckii
CHR	NC 004722	B. cereus	Bacilli	1.00	CHR	NC 010610	L. delordeckii
CHB	NC 006274	B. cereus	Bacilli	1.00	CHB	NC 017465	L fermentum
PL	NC 007107	B. cereus	Bacilli	1.00	CHR	NC_015930	L. garvieae
PL	NC_011341	B. cereus	Bacilli	1.00	CHR	NC_017490	L. garvieae
CHR	NC_011658	B. cereus	Bacilli	1.00	CHR	NC_014319	L. gasicomitatum
CHR	NC_011725	B. cereus	Bacilli	1.00	CHR	NC_{008530}	L. gasseri
CHR	NC_011772	B. cereus	Bacilli	1.00	CHR	NC_018631	L. gelidum
CHR	NC_011773	B. cereus	Bacilli	1.00	CHR	NC_010080	L. helveticus
CHR	NC_011969	B. cereus	Bacilli	1.00	CHR	NC_017467	L. helveticus
CHR	NC_012472	B. cereus	Bacilli	1.00	CHR	NC_018528	L. helveticus
CHR	NC_014335	B. cereus	Bacilli	1.00	CHR	NC 016011	L. innocua
CHR	NC 016779	B. cereus	Bacilli	1.00	CHR	NC 005362	L. iohnsonii
PL	NC 016794	B cereus	Bacilli	0.06	CHR	NC 013504	L. johnsonii
CHR	NC_018491	B. cereus	Bacilli	1.00	CHR	NC_017477	L. johnsonii
CHR	NC_006582	B. clausii	Bacilli	1.00	CHR	NC_015602	L. kefiranofaciens
CHR	NC_015634	B. coagulans	Bacilli	1.00	CHR	NC_014136	L. kimchii
CHR	NC_016023	B. coagulans	Bacilli	1.00	CHR	$NC_{-}002662$	L. lactis
CHR	NC_009674	B. cytotoxicus	Bacilli	1.00	CHR	NC_008527	L. lactis
CHR	NC_002570	B. halodurans	Bacilli	1.00	CHR	NC_009004	L. lactis
CHR	NC_006270	B. licheniformis	Bacilli	1.00	CHR	NC_013656	
CHR	NC_006322	B. licheniformis	Bacilli	1.00	CHR	NC 017486	L. lactis
CHB	NC 014019	B. megaterium	Bacilli	1.00	CHB	NC 017949	L lactis
CHR	NC_014103	B. megaterium	Bacilli	1.00	CHR	NC_008531	L. mesenteroides
CHR	NC_017138	B. megaterium	Bacilli	1.00	CHR	NC_016805	L. mesenteroides
CHR	NC_013791	B. pseudofirmus	Bacilli	1.00	CHR	NC_002973	L. monocytogenes
\mathbf{PL}	NC_013792	B. pseudofirmus	Bacilli	1.00	CHR	NC_003210	L. monocytogenes
CHR	NC_009848	B. pumilus	Bacilli	1.00	CHR	NC_011660	L. monocytogenes
CHR	NC_014219	B. selenitireducens	Bacilli	1.00	CHR	NC_012488	L. monocytogenes
CHR	NC_000964	B. subtilis	Bacilli	1.00	CHR	NC 012768	L. monocytogenes
CHR	NC 014479	B. subtilis	Bacilli	1.00	CHR	NC 017529	L. monocytogenes
PL	NC 015149	B. subtilis	Bacilli	1.00	CHR	NC_017537	L. monocytogenes
CHR	NC_016047	B. subtilis	Bacilli	1.00	CHR	NC_017544	L. monocytogenes
CHR	NC_017195	B. subtilis	Bacilli	1.00	CHR	NC_017545	L. monocytogenes
CHR	NC_018520	B. subtilis	Bacilli	1.00	CHR	NC_017546	L. monocytogenes
CHR	NC_{005957}	B. thuringiensis	Bacilli	1.00	CHR	NC_{017547}	L. monocytogenes
\mathbf{PL}	NC_007203	B. thuringiensis	Bacilli	1.00	CHR	NC_017728	L. monocytogenes
CHR	NC_008600	B. thuringiensis	Bacilli	1.00	CHR	NC_018584	L. monocytogenes
PL	NC_010281	B. thuringiensis	Bacilli	0.86	CHR	NC_018585	L. monocytogenes
СНР	NC 017200	D. mannguensis	Bacilli	1.00	Сив	NC 018587	L monocutogenes
CHB	NC 017208	B thuringiensis	Bacilli	1.00	CHR	NC 018588	L. monocytogenes
CHR	NC_018500	B. thuringiensis	Bacilli	1.00	CHR	NC_018589	L. monocytogenes
CHR	NC_018693	B. thuringiensis	Bacilli	1.00	CHR	NC_018590	L. monocytogenes
CHR	NC_018877	B. thuringiensis	Bacilli	1.00	CHR	NC_018591	L. monocytogenes
CHR	NC_010184	B. weihenstephanensis	Bacilli	1.00	CHR	NC_{018592}	L. monocytogenes
CHR	NC_017743	Bacillus sp.	Bacilli	1.00	CHR	NC_018593	L. monocytogenes
CHR	NC_019425	C. maltaromaticum	Bacilli	1.00	CHR	NC_018642	L. monocytogenes
CHR	NC_015391	Carnobacterium sp.	Bacilli	1.00	CHR	NC_004567	L. plantarum
CHR	NC 004669	E. antarcticum E. faecalic	Bacilli	1.00	CHR	NC 014554	L. plantarum
PI.	NC 004608	E faecalis	Bacilli	0.86	CHR	NC-009513	L. reuteri
CHR	NC_017316	E. faecalis	Bacilli	1.00	CHR	NC_010609	L. reuteri
CHR	NC_018221	E. faecalis	Bacilli	1.00	CHR	NC_015697	L. reuteri
CHR	NC_017022	E. faecium	Bacilli	1.00	CHR	NC_013198	L. rhamnosus
CHR	NC_017960	E. faecium	Bacilli	1.00	CHR	NC_013199	L. rhamnosus
CHR	NC_018081	E. hirae	Bacilli	1.00	CHR	NC_017482	L. rhamnosus
CHR	NC_010556	E. sibiricum	Bacilli	1.00	CHR	NC_017491	L. rhamnosus
CHR	NC_012673	Exiguobacterium sp.	Bacılli Da silli	1.00	CHR	NC_015975	L. ruminis
онк рт	NC 012700	G. kaustopnilus Geobacillus an	Bacilli	nan 1.00	Сир	NC 007020	L. salinarine
CHB	NC 012790	Geobacillus sp.	Bacilli	1.00	CHR	NC.017481	L. salivarius
CHR	NC_013411	Geobacillus sp.	Bacilli	1.00	CHR	NC_015978	L. sanfranciscensis
CHR	NC_014206	Geobacillus sp.	Bacilli	1.00	CHR	NC_013891	L. seeligeri
CHR	NC_014650	Geobacillus sp.	Bacilli	1.00	CHR	NC_015734	Leuconostoc sp.
CHR	NC_014915	Geobacillus sp.	Bacilli	1.00			

CHR	NC_010382	L. sphaericus	Bacilli	1.00	CHR	NC_014494	S. pneumoniae	Bacilli	1.00
CHR	$NC_{-}008555$	L. welshimeri	Bacilli	1.00	CHR	NC_014498	S. pneumoniae	Bacilli	1.00
CHR	NC_011999	M. caseolyticus	Bacilli	1.00	CHR	NC_017591	S. pneumoniae	Bacilli	1.00
CHR	NC_015516	M. plutonius	Bacilli	nan	CHR	NC_017592	S. pneumoniae	Bacilli	1.00
CHR	NC_016938	M. plutonius	Bacilli	1.00	CHR	NC_017593	S. pneumoniae	Bacilli	1.00
CHB	NC 004193	O ihevensis	Bacilli	1.00	CHB	NC 017769	S. prieumoniae	Bacilli	1.00
CHD	NC 008528	O. integenais	Daciili	1.00	CHB	NC 018504	S. pheamoniae	Dacilli	1.00
CUP	NC 016605	P. alguaganii	Dacilli	1.00	CHP	NC 018620	S. pheamoniae	Dacilli	1.00
CHD	NC 008525	D ====t=======	Daciili	1.00	CHR	NC 014025	S. pneumoniae	Dacilli	1.00
CHR	NC-008525	P. pentosaceus	Bacilli	1.00	CHR	NC_014925	S. pseudintermedius	Bacilli	1.00
PL	NC_014628	P. polymyxa	Bacilli	0.17	CHR	NC_017568	S. pseudintermedius	Bacilli	1.00
CHR	NC_004116	S. agalactiae	Bacilli	nan	CHR	NC_015875	S. pseudopneumoniae	Bacilli	1.00
CHR	NC_004368	S. agalactiae	Bacilli	nan	CHR	NC_002737	S. pyogenes	Bacilli	1.00
CHR	NC_007432	S. agalactiae	Bacilli	1.00	CHR	NC_003485	S. pyogenes	Bacilli	1.00
CHR	NC_018646	S. agalactiae	Bacilli	1.00	CHR	NC_004070	S. pyogenes	Bacilli	1.00
CHR	NC_019048	S. agalactiae	Bacilli	1.00	CHR	NC_004606	S. pyogenes	Bacilli	1.00
CHR	NC_002745	S. aureus	Bacilli	1.00	CHR	NC_006086	S. pyogenes	Bacilli	1.00
CHR	NC_002758	S. aureus	Bacilli	1.00	CHR	NC_007296	S. puogenes	Bacilli	1.00
CHB	NC 002951	S aureus	Bacilli	1.00	CHB	NC 007297	S puogenes	Bacilli	1.00
CHR	NC 002952	S aureus	Bacilli	1.00	CHB	NC 008021	S. pyogenes	Bacilli	nan
CHD	NC 002052	C	Daciili	1.00	CHB	NC 008022	S. pybyenes	Dacilli	1 00
CIIL	NC 002955	S. aureus	Dacilli	1.00	CHR	NC_008022	S. pyogenes	Daciili	1.00
CIIL	NC_005925	S. aureus	D Saciili	1.00	CHR	NC_008023	S. pyogenes	Daciili	1.00
CHR	NC-007622	S. aureus	Bacilli	1.00	CHR	NC_008024	S. pyogenes	Bacilli	1.00
CHR	NC_007793	S. aureus	Bacilli	1.00	CHR	NC_009332	S. pyogenes	Bacilli	1.00
CHR	NC_007795	S. aureus	Bacilli	1.00	CHR	NC_011375	S. pyogenes	Bacilli	1.00
CHR	NC_009487	S. aureus	Bacilli	1.00	CHR	NC_017040	S. pyogenes	Bacilli	1.00
CHR	NC_009632	S. aureus	Bacilli	1.00	CHR	NC_017053	S. pyogenes	Bacilli	1.00
CHR	NC_009641	S. aureus	Bacilli	1.00	CHR	NC_017596	S. pyogenes	Bacilli	1.00
CHR	NC_009782	S. aureus	Bacilli	1.00	CHR	NC_018936	S. pyogenes	Bacilli	1.00
CHR	NC_010079	S. aureus	Bacilli	1.00	CHR	NC_015760	S. salivarius	Bacilli	1.00
PL	NC_013329	S. aureus	Bacilli	0.73	CHR	NC_017595	S. salivarius	Bacilli	1.00
PL	NC_013341	S. aureus	Bacilli	0.53	CHB	NC_009009	S. sanquinis	Bacilli	1.00
CHR	NC_013450	S. aureus	Bacilli	1.00	CHR	NC 007350	S. sapronhuticus	Bacilli	1.00
CHP	NC 016012	S aureus	Bacilli	1.00	СИР	NC 018065	S silmestrie	Bacilli	1.00
CUP	NC 016028	S. dureus	Dacilli	1.00	DI	NC 018060	S. silvestris	Dacilli	0.86
СПБ	NC 016041	S aureas	Da.:11:	1.00		NC 000449	S. SHUCSHIS S. SHUCSHIS	Pacilli	1.00
CIIL	NC 017221	S. aureus	Dacilli	1.00	CHR	NC_009442	S. suis	Daciili	1.00
CHR	NC_017331	S. aureus	Bacilli	1.00	CHR	NC_009443	S. suis	Bacilli	1.00
CHR	NC_017333	S. aureus	Bacilli	1.00	CHR	NC_012924	S. suis	Bacilli	1.00
CHR	NC_017337	S. aureus	Bacilli	1.00	CHR	NC_012925	S. suis	Bacilli	1.00
CHR	NC_017338	S. aureus	Bacilli	1.00	CHR	NC_012926	S. suis	Bacilli	1.00
CHR	NC_017340	S. aureus	Bacilli	1.00	CHR	NC_015433	S. suis	Bacilli	1.00
CHR	NC_017341	S. aureus	Bacilli	1.00	CHR	NC_017617	S. suis	Bacilli	1.00
CHR	NC_017342	S. aureus	Bacilli	1.00	CHR	NC_017618	S. suis	Bacilli	1.00
CHR	NC_017343	S. aureus	Bacilli	1.00	CHR	NC_017619	S. suis	Bacilli	1.00
CHR	NC_017347	S. aureus	Bacilli	1.00	CHR	NC_017620	S. suis	Bacilli	1.00
CHR	NC_017349	S. aureus	Bacilli	1.00	CHR	NC_017621	S. suis	Bacilli	1.00
CHR	NC_017351	S. aureus	Bacilli	1.00	CHR	$NC_{-}017622$	S. suis	Bacilli	1.00
CHR	NC 017673	S. aureus	Bacilli	1.00	CHR	NC 017950	S. suis	Bacilli	1.00
CHB	NC 017763	S aureus	Bacilli	1.00	CHB	NC 018526	S suis	Bacilli	1.00
CHR	NC 018608	S. aureus	Bacilli	1.00	CHB	NC 006448	S. thermonbilus	Bacilli	1.00
CUP	NC 012121	S. dureus	Dacilli	1.00	CHP	NC 006440	S. thermorphilus	Dacilli	1.00
CHD	NC 012021	S. curnosus	Daciili	1.00	CHR	NC 008522	S. thermophilus	Dacilli	1.00
CIIL	NC_012691	S. aysgatactiae	D Saciili	1.00	CHR	NC_008552	S. inermophilus	Daciili	1.00
CHR	NC_017567	S. dysgalactiae	Bacilli	1.00	CHR	NC_017563	S. thermophilus	Bacilli	1.00
CHR	NC_018712	S. dysgalactiae	Bacıllı	1.00	CHR	NC_017581	S. thermophilus	Bacıllı	1.00
CHR	NC_019042	S. dysgalactiae	Bacilli	1.00	CHR	NC_017927	S. thermophilus	Bacilli	1.00
CHR	NC_002976	S. epidermidis	Bacilli	1.00	CHR	NC_012004	S. uberis	Bacilli	1.00
CHR	NC_004461	S. epidermidis	Bacilli	1.00	CHR	NC_016052	T. halophilus	Bacilli	1.00
CHR	NC_011134	S. equi	Bacilli	1.00	CHR	NC_015759	W. koreensis	Bacilli	1.00
CHR	NC_012470	S. equi	Bacilli	1.00	CHR	NC_014378	A. arabaticum	Clostridia	1.00
CHR	NC_012471	S. equi	Bacilli	1.00	CHR	NC_013385	A. degensii	Clostridia	1.00
CHR	NC_017582	S. equi	Bacilli	1.00	PL	NC_013386	A. degensii	Clostridia	0.91
CHR	NC_013798	S. gallolyticus	Bacilli	1.00	CHR	NC_009633	A. metalliredigens	Clostridia	1.00
CHR	NC_015215	S. galloluticus	Bacilli	1.00	CHR	NC_009922	A. oremlandii	Clostridia	1.00
CHR	NC_017576	S. aalloluticus	Bacilli	1.00	PL	NC_013164	A. prevotii	Clostridia	1.00
CHB	NC 009785	S aordonii	Bacilli	1.00	CHB	NC 013171	A prevotii	Clostridia	1.00
CHB	NC 007168	S haemoluticus	Bacilli	1.00	CHB	NC 016894	A woodii	Clostridia	1.00
CHR	NC 016826	S. infantarius	Bacilli	1.00	CHR	NC 014387	B. proteoclasticus	Clostridia	1.00
CHR	NC 018073	S intermediae	Bacilli	1.00	СНВ	NC 003030	C acetobatalicam	Clostridia	1.00
CHB	NC 013803	S. Indunensis	Bacilli	1.00	СПК	NC 015687	C. acetobutulicum	Clostridia	1.00
СПБ	NC 017959	S lugademonoi-	Daciiii Daciiii	1.00		NC 017905	C acetobatali	Clost	1.00
CUD	NC 016740	S. ruguunensis	Daciiii D:11:	1.00	CHR	NC 019664	C. acidumici	Clostridia	1.00
CUT	NC 0100749	S. maceaonicus	Daciiii	1.00	CHR	NC 015016		Clostridia	1.00
CHR	NC-013853	S. mitis	Bacilli	1.00	CHR	NC_015913	Cana. Arthromitus	Clostridia	1.00
CHR	NC_004350	S. mutans	Bacilli	1.00	CHR	NC_016012	Cand. Arthromitus	Clostridia	1.00
CHR	NC_013928	5. mutans	Bacilli	1.00	CHR	NG_017294	Cana. Arthromitus	Clostridia	1.00
CHR	NC_017768	S. mutans	Bacilli	1.00	CHR	NC_009617	C. beijerinckii	Clostridia	1.00
CHR	NC_018089	S. mutans	Bacilli	1.00	CHR	NC_009495	C. botulinum	Clostridia	1.00
CHR	NC_015291	S. oralis	Bacilli	1.00	PL	NC_009496	C. botulinum	Clostridia	0.79
CHR	NC_{015678}	$S. \ parasanguinis$	Bacilli	1.00	CHR	$NC_{-}009697$	C. botulinum	Clostridia	1.00
CHR	NC_017905	$S. \ parasanguinis$	Bacilli	1.00	CHR	NC_009698	C. botulinum	Clostridia	1.00
CHR	NC_{015558}	S. parauberis	Bacilli	1.00	CHR	NC_009699	C. botulinum	Clostridia	1.00
PL	NC_007167	S. pasteuri	Bacilli	1.00	CHR	NC_010516	C. botulinum	Clostridia	1.00
CHR	NC_015600	S. pasteurianus	Bacilli	1.00	CHR	NC_010520	C. botulinum	Clostridia	1.00
CHR	NC_003028	S. pneumoniae	Bacilli	nan	CHR	NC_010674	C. botulinum	Clostridia	1.00
CHB	NC 003098	S nneumoniae	Bacilli	1 00	СПР	NC 010722	C hotulinum	Clostridia	1.00
СПР	NC 008522	S. pricamoniae	Daciiii Daciiii	1.00		NC 010723	C hotalinam	Clostridia	1.00
CITE	NC 010200	S. pneumoniae	Bacilli	1.00	CHR	NC 012563	C. botulinum	Clostridia	1.00
CHR	NC 010500	5. pneumoniae	Bacilli	1.00	CHR	NC 012058	C. botuinum	Clostridia	1.00
CHR	NC_010582	5. pneumoniae	Bacilli	1.00	PL	NG_012945	C. botulinum	Clostridia	0.81
CHR	NC_011072	S. pneumoniae	Bacilli	1.00	PL	NC_015417	C. botulinum	Clostridia	1.00
CHR	NC_011900	S. pneumoniae	Bacilli	1.00	PL	NC_015418	C. botulinum	Clostridia	1.00
CHR	NC_012466	$S. \ pneumoniae$	Bacilli	1.00	CHR	$NC_{-}015425$	C. botulinum	Clostridia	1.00
CHR	NC_{012467}	$S. \ pneumoniae$	Bacilli	1.00	PL	$NC_{-}015426$	C. botulinum	Clostridia	0.71
CHR	NC_012468	$S. \ pneumoniae$	Bacilli	1.00	CHR	NC_017297	C. botulinum	Clostridia	1.00
CHR	NC_012469	S. pneumoniae	Bacilli	1.00	CHR	NC_017299	C. botulinum	Clostridia	1.00
CHR	NC_014251	S. pneumoniae	Bacilli	1.00	CHR	NC_011898	C. cellulolyticum	Clostridia	1.00
-	-	 ************************************		-			9		

	NC 014393	C cellulovorans	Clostridia	1 00	CHB	NC 013203	A parvulum	Coriobacteridae	0.67
CHR	NC_016627	C. clariflavum	Clostridia	0.88	CHR	NC_013170	C. curtum	Coriobacteridae	1.00
CHR	NC_010424	Cand. Desulforudis	Clostridia	1.00	CHR	NC_015389	C. alomerans	Coriobacteridae	1.00
CHR	NC_009089	C. difficile	Clostridia	nan	CHR	NC_013204	E. lenta	Coriobacteridae	1.00
CHR	NC_013315	C. difficile	Clostridia	0.50	CHR	NC_015738	Eggerthella sp.	Coriobacteridae	1.00
CHR	NC_013316	C. difficile	Clostridia	0.43	CHR	NC_014363	O. uli	Coriobacteridae	1.00
CHR	NC_017179	C. difficile	Clostridia	0.47	CHR	NC_013165	S. heliotrinireducens	Coriobacteridae	1.00
CHR	NC_013895	Clostridiales sp.	Clostridia	1.00	\mathbf{PL}	NC_013940	D. desulfuricans	Deferribacteres	0.50
CHR	NC_007503	C. hydrogenoformans	Clostridia	1.00	CHR	NS_000195	C. Cloacamonas	Division WWE1	1.00
CHR	NC_009706	C. kluyveri	Clostridia	1.00	CHR	NC_010644	E. minutum	Elusimicrobia	1.00
CHR	NC_011837	C. kluyveri	Clostridia	1.00	CHR	NS_000191	Termite group 1	Elusimicrobia	0.68
PL	NC_014719	C. kristjanssonii	Clostridia	1.00	CHR	NC_015601	E. rhusiopathiae	Erysipelotrichia	1.00
CHR	NC_015275	C. lentocellum	Clostridia	1.00	CHR	NC_003454	F. nucleatum	Fusobacteriia	1.00
CHR	NC_014328	C. ljungdahlii	Clostridia	1.00	CHR	NC_014632	I. polytropus	Fusobacteriia	1.00
CHR	NC_008593	C. novyi	Clostridia	1.00	RECE	NC_014633	I. polytropus	Fusobacteriia	1.00
CHR	NC_003366	C. perfringens	Clostridia	1.00	CHR	NC_013192	L. buccalis	Fusobacteriia	1.00
CHR	NC_008261	C. perfringens	Clostridia	1.00	CHR	NC_013515	S. moniliformis	Fusobacteriia	0.86
CHR	NC_008262	C. perfringens	Clostridia	1.00	CHR	NC_013517	S. termitidis	Fusobacteriia	0.67
CHR	NC_010001	C. phytofermentans	Clostridia	1.00	CHR	NC_010163	A. laidlawii	Mollicutes	1.00
CHR	NC_014376	C. saccharolyticum	Clostridia	1.00	CHR	NC_011047	Cand. Phytoplasma	Mollicutes	1.00
CHR	NC_015737	Clostridium sp.	Clostridia	1.00	CHR	NC_011025	M. arthritidis	Mollicutes	1.00
CHR	NC_016791	Clostridium sp.	Clostridia	0.00	CHR	NC_014014	M. crocodyli	Mollicutes	1.00
CHR	NC_014614	C. sticklandii	Clostridia	1.00	CHR	NC_006055	M. florum	Mollicutes	1.00
CHR	NC_004557	C. tetani	Clostridia	1.00	CHR	NC_013511	M. hominis	Mollicutes	1.00
PL	NC_004565	C. tetani	Clostridia	0.00	CHR	NC_006908	M. mobile	Mollicutes	1.00
CHR	NC_013216	D. acetoxidans	Clostridia	1.00	CHR	NC_002771	M. pulmonis	Mollicutes	1.00
CHR	NC_018068	D. acidiphilus	Clostridia	1.00	CHR	NC_007294	M. synoviae	Mollicutes	1.00
CHR	NC_015565	D. carboxydivorans	Clostridia	1.00	CHR	NC_017080	P. mikurensis	Phycisphaerae	0.96
CHR	NC_018017	D. dehalogenans	Clostridia	1.00	\mathbf{PL}	NC_014149	P. limnophilus	Planctomycetacia	1.00
CHR	NC_007907	D. hafniense	Clostridia	0.73	CHR	NC_014962	I. pallida	Planctomycetacia	0.14
CHR	NC_011830	D. hafniense	Clostridia	1.00	\mathbf{PL}	NZ_CM000773	3 Rickettsia endosymbiont	α proteobacteria	0.86
CHR	NC_015573	D. kuznetsovii	Clostridia	nan	CHR	NC_016620	B. marinus	δ proteobacteria	nan
CHR	NC_018515	D. meridiei	Clostridia	1.00	CHR	NC_009943	D. oleovorans	δ proteobacteria	0.71
CHR	NC_016584	D. orientis	Clostridia	1.00	\mathbf{PL}	NC_012797	D. magneticus	δ proteobacteria	0.64
CHR	NC_009253	D. reducens	Clostridia	1.00	\mathbf{PL}	NC_008014	L. intracellularis	δ proteobacteria	0.79
CHR	NC_015589	D. ruminis	Clostridia	1.00	\mathbf{PL}	NC_005915	C. taeniospiralis	γ proteobacteria	0.77
CHR	NC_018866	Dehalobacter sp.	Clostridia	1.00	CHR	NC_002967	T. denticola	Spirochaetes	0.12
CHR	NC_018867	Dehalobacter sp.	Clostridia	1.00	CHR	NC_014011	A. colombiense	Synergistia	1.00
CHR	NC_012778	E. eligens	Clostridia	1.00	CHR	NC_018024	A. mobile	Synergistia	1.00
PL	NC_012780	E. eligens	Clostridia	1.00	CHR	NC_013522	$T. \ acidaminovorans$	Synergistia	1.00
PL	NC_012782	E. eligens	Clostridia	0.58	CHR	NC_016148	T. lienii	Synergistia	1.00
CHR	NC_014828	E. harbinense	Clostridia	1.00	CHR	NC_012785	K. olearia	Thermotogae	0.23
CHR	NC_014624	E. limosum	Clostridia	nan	CHR	NC_017934	M. prima	Thermotogae	1.00
CHR	NC_012781	E. rectale	Clostridia	1.00					
CHR	NC_016630	F. alocis	Clostridia	1.00	cluster :	2 stability mea	sure : 0.817155		
PL	NC_010371	F. magna	Clostridia	1.00	CHR	NC_014497	Cand. Zinderia	β proteobacteria	1.00
CHR	NC_010376	F. magna	Clostridia	1.00	CHR	NC_016513	A.actinomycetem comitans	γ proteobacteria	1.00
CHR	NC_014654	H. hydrogeniformans	Clostridia	1.00	CHR	NC_017846	A.actinomycetem comitans	γ proteobacteria	1.00
CHR	NC 010337	II and a death and dean	C11 11.	1 1 00	OIID		4 7 .		
=-	100_010357	n. moaesticataum	Clostridia	1.00	CHR	NC_009053	A. pleuropneumoniae	γ proteobacteria	1.00
CHR	NC_011899	H. orenii	Clostridia	1.00	CHR	NC_009053 NC_010278	A. pleuropneumoniae A. pleuropneumoniae	γ proteobacteria γ proteobacteria	1.00 1.00
CHR CHR	NC_011899 NC_017455	H. moaesticataam H. orenii H. praevalens	Clostridia Clostridia Clostridia	1.00 1.00 1.00	CHR CHR CHR	NC_009053 NC_010278 NC_010939	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae	γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00
CHR CHR CHR	NC_011899 NC_017455 NC_015520	H. moaesticataum H. orenii H. praevalens M. australiensis	Clostridia Clostridia Clostridia	1.00 1.00 1.00	CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00 1.00
CHR CHR CHR CHR	NC_011899 NC_017455 NC_015520 NC_007644	H. modesticalitam H. orenii H. praevalens M. australiensis M. thermoacetica	Clostridia Clostridia Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00	CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00 1.00 1.00
CHR CHR CHR CHR CHR	NC_011899 NC_017455 NC_015520 NC_007644 NC_010718	H. modesitataan H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus	Clostridia Clostridia Clostridia Clostridia Clostridia	$ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 $	CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$ \begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array} $
CHR CHR CHR CHR CHR PL	NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010724	 n. modestication H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00 0.86	CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida A. salmonicida	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array} $
CHR CHR CHR CHR CHR PL PL	NC_011939 NC_017455 NC_017455 NC_007644 NC_010718 NC_010724 NC_016046	H. modesitediatum H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus O. valericigenes	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ \end{array} $	CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida A. salmonicida A. suis	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array} $
CHR CHR CHR CHR CHR PL PL CHR	NC_011939 NC_017455 NC_015520 NC_007644 NC_010718 NC_010724 NC_016046 NC_016048	 h. modesitediatum H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus O. valericigenes O. valericigenes 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida A. salmonicida A. suis A. veronii	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
CHR CHR CHR CHR PL PL CHR CHR CHR	NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010724 NC_016046 NC_016048	 n. modestication H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR	NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010724 NC_016046 NC_016048 NC_009454 NC_009454	 n. modesstradium H. orenti H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
CHR CHR CHR CHR PL CHR CHR CHR CHR CHR	NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_016046 NC_016048 NC_009454 NC_014833 NC_015977	 h. modesticatuum H. orenii H. praevalens M. australiensis M. thermoacctica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR PL CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011138	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii	γ proteobacteria γ proteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
CHR CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR	NC_011899 NC_017455 NC_015520 NC_007644 NC_010724 NC_010724 NC_016046 NC_009454 NC_014833 NC_015977 NC_015172	 h. modesticadaum H. orenii H. praevalens M. australiensis M. thermoactica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00 0.86 0.62 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR CHR CHR CHR CHR CHR CHR CHR CHR PL CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011138 NC_018632	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. sulmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii	γ proteobacteria γ proteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR	NC_011899 NC_017455 NC_015520 NC_007644 NC_010724 NC_010724 NC_016046 NC_009454 NC_014833 NC_015977 NC_015172 NC_014220	 n. modestication H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus S. glycolicus S. lipocalidus 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_018653 NC_011138 NC_0118632 NC_018678	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii	γ proteobacteria γ proteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_009454 NC_014833 NC_015977 NC_015172 NC_014220 NC_016040	 n. modescitation H. orenti H. praevalens M. australiensiss M. thermoacetica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.21\\ 1.02\\ 0.21\\ 1.02\\ 0.21\\ 1.02\\ 0.22\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011138 NC_018675 NC_018677 NC_018679	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salcodii A. macleodii A. macleodii	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_009454 NC_009454 NC_009454 NC_014833 NC_015977 NC_014220 NC_014220 NC_016040 NC_008346	 h. modestication H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei 	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.21\\ 1.00\\ 0.22\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_009355 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_01138 NC_018653 NC_018679 NC_018679 NC_018692	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. suis A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_015977 NC_014220 NC_016040 NC_008346 NC_015519	 n. modestication H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.21\\ 1.00\\ 1.00\\ 0.21\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_01138 NC_016853 NC_01138 NC_018679 NC_018679 NC_018692 NC_019393 NC_0155	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. suis A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii	Pproteobacteria pproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010724 NC_010724 NC_010646 NC_009454 NC_014833 NC_015977 NC_014220 NC_01640 NC_008346 NC_015519 NC_014964	 n. modestication H. orenii H. praevalens M. australiensis M. thermocctica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. itsi urg 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_009555 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_018653 NC_011138 NC_018632 NC_018678 NC_018679 NC_018692 NC_018554	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. succinode A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A.	Pproteobacteria pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010718 NC_010714 NC_010724 NC_016046 NC_010483 NC_009454 NC_014833 NC_015472 NC_015172 NC_015172 NC_015172 NC_015179 NC_016040 NC_008346 NC_015519 NC_014964 NC_013921	 n. modesticulariti H. orenii H. praevalens M. australiensiss M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus P. thermopropionicum R. albus S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. titalicus T. more service 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424 NC_018690 NC_016853 NC_011138 NC_018632 NC_018632 NC_018679 NC_018679 NC_018679 NC_018554 NC_002528	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. theromonas sp. B. aphidicola	Pproteobacteria pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_015977 NC_014220 NC_014020 NC_014964 NC_013921 NC_014831 NC_014831	 n. modestication H. oraenii H. praevalens M. australiensis M. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_009348 NC_011312 NC_018690 NC_015424 NC_018633 NC_016853 NC_018633 NC_018679 NC_018679 NC_018679 NC_018679 NC_015554 NC_002528 NC_004645	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. suis A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola	Pproteobacteria Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_010724 NC_014833 NC_014833 NC_014833 NC_014833 NC_014977 NC_014200 NC_016040 NC_008346 NC_013921 NC_014831 NC_014209	 n. modestication H. orenii H. praevalens M. australiensis M. thermoperities N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matinensis T. mathranii 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_008570 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_01138 NC_016853 NC_01138 NC_018679 NC_018679 NC_018679 NC_018679 NC_018675 NC_012528 NC_004061 NC_004545	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. suis A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola	Pproteobacteria Pproteobacteria	1.00 1.00
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010718 NC_010718 NC_0106046 NC_016048 NC_009454 NC_016048 NC_015977 NC_015172 NC_015172 NC_015172 NC_016040 NC_008346 NC_015519 NC_014964 NC_013921 NC_014209 NC_014209 NC_014209 NC_014209	 n. modestication H. modestication H. praevalens M. australiensis M. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. thermotolerans S. thermotolerans S. thermotolerans T. acetatoxydans T. brockii T. italicus T. marianensis T. narugense 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_00955 NC_019312 NC_018690 NC_015424 NC_012913 NC_018633 NC_011138 NC_018632 NC_018678 NC_018679 NC_018692 NC_018692 NC_018554 NC_002528 NC_004061 NC_004545 NC_008513	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. succinode A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola B. aphidicola	Pproteobacteria Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010718 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_014977 NC_015172 NC_014220 NC_014200 NC_014920 NC_014964 NC_014209 NC_014209 NC_014209 NC_014209 NC_014209 NC_014209 NC_014277	 n. modesticadadm H. orenti H. praevalens M. australiensiss M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus P. thermopropionicum R. albus S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. marianensis T. mariagense T. oceani T. brockii 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_009348 NC_011312 NC_015424 NC_012913 NC_016853 NC_011453 NC_018673 NC_018679 NC_018679 NC_018679 NC_018679 NC_01854 NC_002528 NC_004545 NC_004545 NC_004513 NC_011833 NC_011833	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola B. aphidicola B. aphidicola	Pproteobacteria Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_015977 NC_014220 NC_014220 NC_014040 NC_013921 NC_014831 NC_014209 NC_015499 NC_015499 NC_015499 NC_015497 NC_018870	 n. modestication H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matinanensis T. matyanensis T. oceani T. phaeum T. phaeum 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_009355 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011318 NC_018672 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_015554 NC_002528 NC_004513 NC_004513 NC_011833 NC_011833 NC_011834	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola B. aphidicola B. aphidicola	Pproteobacteria Ppr	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010718 NC_010724 NC_016046 NC_009454 NC_014833 NC_015977 NC_014833 NC_015497 NC_016040 NC_008346 NC_013921 NC_014831 NC_014831 NC_014831 NC_014831 NC_014831 NC_014831 NC_014831 NC_014831 NC_014831 NC_014831 NC_014837 NC_014837 NC_014870 NC_014870 NC_014870 NC_014870 NC_014870	 n. modestication H. modestication H. praevalens M. australiensis M. thermopatiensis N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matianensis T. mathranii T. narugense T. oceani T. phage T. ratera 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_009357 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_01138 NC_016853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018675 NC_004545 NC_002528 NC_004061 NC_004545 NC_008513 NC_011833 NC_011834 NC_015662	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola	Pproteobacteria Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010718 NC_010718 NC_010718 NC_0106046 NC_016048 NC_016048 NC_01519 NC_015172 NC_015172 NC_015172 NC_015172 NC_015179 NC_015519 NC_014964 NC_013921 NC_014831 NC_014209 NC_014831 NC_014209 NC_014377 NC_018870 NC_017992 NC_014152	 n. modestication H. modestication H. praevalens M. australiensis M. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. thermotolerans S. thoremotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. mathranii T. narugense T. oceani T. phaeum T. phage T. potens 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_01138 NC_016853 NC_0118632 NC_018679 NC_018679 NC_018692 NC_018692 NC_018692 NC_002528 NC_004061 NC_004545 NC_004545 NC_004545 NC_001833 NC_011834 NC_015662 NC_017252	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. sus A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_010724 NC_016046 NC_010483 NC_009454 NC_014833 NC_014977 NC_015172 NC_014220 NC_014220 NC_014924 NC_014964 NC_014991 NC_014999 NC_014377 NC_017992 NC_014152 NC_010321 NC_010321	 n. modestication H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. motockii T. marianensis T. mathranii T. narugense T. oceani T. phaeum T. phaeum T. potens T. potens T. potens T. potens 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009555 NC_009348 NC_011312 NC_018690 NC_011312 NC_018633 NC_011313 NC_018633 NC_018632 NC_018678 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_00451554 NC_004513 NC_004513 NC_011833 NC_011833 NC_011834 NC_017252 NC_017252 NC_017252	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. aphrophilus A. fischeri A. macleodii B. aphidicola 	Pproteobacteria Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_015977 NC_015172 NC_014220 NC_014200 NC_014209 NC_014377 NC_015499 NC_014377 NC_018870 NC_014329 NC_014152 NC_010321 NC_010320	 n. modestication H. orenii H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matinanensis T. matinanensis T. maturanii T. narugense T. oceani T. phaeum T. phaeum T. piaeum T. piaeus T. potens T. pseudethanolicus Thermoanaerobacter sp. 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_009555 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011318 NC_016853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_015554 NC_002528 NC_004513 NC_001553 NC_011833 NC_011833 NC_011834 NC_017252 NC_017253 NC_017253	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_014833 NC_015977 NC_014200 NC_014200 NC_014040 NC_013921 NC_014991 NC_014831 NC_014391 NC_014391 NC_014377 NC_014870 NC_014377 NC_018870 NC_014371 NC_014321 NC_010321 NC_010320 NC_014538 NC_00326	 h. modesicialiam H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus O. valericigenes O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matinensis T. matinensis T. maturensis T. maturensis T. phage T. potens T. poseudethanolicus Thermoanaerobacter sp. T. tenacacaranie 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011318 NC_016853 NC_0116853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_004545 NC_002528 NC_004061 NC_004545 NC_002523 NC_011833 NC_011833 NC_011834 NC_017252 NC_017254 NC_017255	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fascheri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010718 NC_010718 NC_010718 NC_010724 NC_016046 NC_016048 NC_014833 NC_015172 NC_015172 NC_015172 NC_015172 NC_015172 NC_015179 NC_01519 NC_014964 NC_013921 NC_014209 NC_014209 NC_014377 NC_014209 NC_014377 NC_014377 NC_014377 NC_014377 NC_014371 NC_014377 NC_014371 NC_014371 NC_014320 NC_014321 NC_010320 NC_014538 NC_003460 NC_014416	 n. modessicalian H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acteatoxydans T. brockii T. italicus T. marianensis T. marianensis T. marianensis T. marianensis T. phaeum T. phaeum T. potens T. peudethanolicus Thermoanaerobacter sp. T. tengcongensis 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\ 0.86\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010278 NC_010939 NC_009555 NC_009555 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_0116853 NC_0116853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018692 NC_018679 NC_018679 NC_018679 NC_018554 NC_004545 NC_004545 NC_004545 NC_004562 NC_017252 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. hydrophila A. salmonicida A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_015977 NC_015172 NC_014220 NC_014220 NC_014220 NC_014202 NC_014964 NC_013921 NC_014964 NC_01499 NC_01499 NC_01499 NC_014377 NC_018370 NC_017992 NC_014152 NC_010320 NC_014538 NC_003869 NC_014410 NC_014555	 n. modestication H. orenii H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. motockii T. noceani T. phaeum T. phaeum T. piseudethanolicus Thermoanaerobacter sp. T. thermoanechi T. thermosaccharolyticum 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009555 NC_009348 NC_011312 NC_018690 NC_011312 NC_018673 NC_011313 NC_018633 NC_011333 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_00451554 NC_004513 NC_004513 NC_011833 NC_011833 NC_017252 NC_017255 NC_017255 NC_017255 NC_017256 NC_017256 NC_017256 NC_017256	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinoda A. salmonicida A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. aphidicola	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_015977 NC_014172 NC_014200 NC_014200 NC_014040 NC_014921 NC_014991 NC_014377 NC_015499 NC_014377 NC_015499 NC_014377 NC_018870 NC_014312 NC_014312 NC_014312 NC_014312 NC_014538 NC_003869 NC_014410 NC_015558	 n. modestication H. orenii H. praevalens M. australiensis M. thermoactica N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matiranensis T. maturanii T. narugense T. oceani T. phage T. potens T. pseudethanolicus Thermoanaerobacter sp. Thermoanaerobacter sp. T. thermoacharolyticum T. wiegelii T. mathrani 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 0.86\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 1.00\\ 1.00\\ 0.86\\ 1.00\\ 1.00\\ 0.86\\ 1.00\\ 0.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011318 NC_016853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_015554 NC_002528 NC_004513 NC_011833 NC_011833 NC_011833 NC_011833 NC_011833 NC_017252 NC_017255 NC_017255 NC_017255 NC_017256 NC_017259 NC_00784	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. exprophilus A. aphrophilus A. fischeri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_014833 NC_015977 NC_014833 NC_015977 NC_014200 NC_014040 NC_008346 NC_013921 NC_014964 NC_013921 NC_014377 NC_014831 NC_014399 NC_014377 NC_014870 NC_014317 NC_014312 NC_014321 NC_010320 NC_014538 NC_003869 NC_014410 NC_015558 NC_015558	 n. modestication H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus O. valericigenes O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. mathranii T. narugense T. optens T. pseudethanolicus Thermoanaerobacter sp. T. tengcongensis T. thermoanaerobacter T. tengcongensis T. tuigelii T. xylanolyticum A fermoataerotacies 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011318 NC_016853 NC_0116853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_002528 NC_004061 NC_004545 NC_002528 NC_017255 NC_017254 NC_017255 NC_017255 NC_017255 NC_017256 NC_017256 NC_007984 NC_005061	A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. succinogenes A. sus A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fascheri A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii A. macleodii B. aphidicola B. a	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010718 NC_010718 NC_010724 NC_016046 NC_010646 NC_016048 NC_015172 NC_015172 NC_015172 NC_015172 NC_015172 NC_015172 NC_015179 NC_015519 NC_014964 NC_013921 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_017992 NC_014152 NC_010320 NC_014538 NC_003869 NC_014555 NC_015555 NC_015555 NC_013740	 n. modestication H. orenti H. orenti H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. marianensis T. marianensis T. narugense T. oceani T. phaeum T. phage T. potens Thermoanaerobacter sp. Thermoanaerobacter sp. T. tengcongensis T. thermosaccharolyticum T. wigelii T. xylanolyticum A. fermentans 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_009348 NC_011312 NC_015424 NC_012913 NC_016853 NC_01138 NC_018673 NC_018673 NC_018678 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018513 NC_01554 NC_004545 NC_004545 NC_004545 NC_004545 NC_004545 NC_017252 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017259 NC_007984 NC_005061 NC_0079292	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. sucinogenes A. sucinogenes A. salmonicida A. macleodii B. aphidicola B. aphidi	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_014977 NC_015172 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_01420 NC_014964 NC_013921 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_014519 NC_017992 NC_014152 NC_010320 NC_014538 NC_003869 NC_014555 NC_015555 NC_013740 NC_016977	 n. modesticularin H. orenii H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. motockii T. norugense T. oceani T. phaeum T. phaeum T. piaeum T. potens T. pseudethanolicus Thermoanaerobacter sp. T. thermosaccharolyticum T. wiegelii T. zylanolyticum A. intestini M. alkderzii 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009555 NC_009348 NC_011312 NC_018650 NC_001312 NC_018679 NC_018673 NC_018673 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_002528 NC_004545 NC_004545 NC_004545 NC_004545 NC_0017252 NC_017252 NC_017252 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017256 NC_017255 NC_017256 NC_017256 NC_007984 NC_007920 NC_007921	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. salmonicida A. suis A. shrophilus A. fischeri A. macleodii B. aphidicola B. aphid	Pproteobacteria Ppr	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_017455 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_0104833 NC_01954 NC_014833 NC_014200 NC_014200 NC_014200 NC_014901 NC_014901 NC_014991 NC_014377 NC_015499 NC_014377 NC_015499 NC_014377 NC_018870 NC_014321 NC_014321 NC_014328 NC_014525 NC_014555 NC_013740 NC_015873 NC_015873	 n. modestication H. orenii H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. matinanensis T. mathranii T. noreugense T. oceani T. phaeum T. phaeum T. phaeum T. piaeum T. potens T. pseudethanolicus Thermoanaerobacter sp. Thermoanaerobacter sp. Thermoanaerobacter sp. T. thermosacharolyticum T. wiegelii T. xylanolyticum A. fermentans A. intestini M. elsdenii S. ruminartium 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_009355 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011312 NC_016853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_015554 NC_002528 NC_004513 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017256 NC_017259 NC_007984 NC_007994 NC_007292 NC_014909 NC_008512	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. suis A. ephrophilus A. fischeri A. macleodii B. aphidicola Cand. Blochmannia Cand. Carsonella Cand. Harvitarvella 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_014833 NC_015977 NC_014833 NC_015977 NC_014200 NC_016040 NC_013921 NC_01499 NC_01499 NC_014311 NC_014311 NC_014311 NC_014321 NC_014321 NC_014312 NC_014321 NC_014321 NC_014321 NC_014321 NC_014321 NC_014321 NC_014321 NC_014321 NC_014538 NC_003869 NC_014410 NC_015555 NC_013740 NC_015873 NC_017620	 n. modestication H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matraensis T. matraensis T. matraense T. optens T. pseudethanolicus Thermoanaerobacter sp. Thermoanaerobacter sp. T. tengcongensis T. thermoanaerobacter sp. T. tengcongensis T. thermoanaerobacter sp. T. tengcongensis T. thermoanaerobacter sp. T. tengcongensis A. intestini M. elsdenii S. ruminantium 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011318 NC_016853 NC_0116853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_002528 NC_004061 NC_002528 NC_0017255 NC_007984 NC_005061 NC_007912 NC_007912 NC_007512 NC_012751 NC_007502	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. sulmonicida A. salmonicida A. salmonicida A. suis A. earnii A. aphrophilus A. macleodii B. aphidicola B. a	Pproteobacteria	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010718 NC_010718 NC_010718 NC_010718 NC_010464 NC_009454 NC_015977 NC_015172 NC_015172 NC_015172 NC_015172 NC_015179 NC_015519 NC_014964 NC_015519 NC_014904 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_014555 NC_013740 NC_015873 NC_017068 NC_017068 NC_017079	 n. modessicalization H. orenti H. orenti H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. marianensis T. marianensis T. marianensis T. phaeum T. phage T. potens T. peudethanolicus Thermoanaerobacter sp. T. tengcongensis T. thermosaccharolyticum T. wiegelii T. xylanolyticum A. fermentans A. intestini S. ruminantium S. ruminantium 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\ 0.88\\ 0 \\ 88\\ 0 \\ 88\\ 0 \\ 88\\ 0 \\ 0 \\ 88\\ 0 \\ 0 $	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009655 NC_009348 NC_011312 NC_015424 NC_012913 NC_016853 NC_011424 NC_012913 NC_016853 NC_012913 NC_018678 NC_018678 NC_018679 NC_018679 NC_018679 NC_018679 NC_019393 NC_01554 NC_004545 NC_004545 NC_004545 NC_004545 NC_004545 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_007984 NC_005061 NC_008512 NC_019721 NC_008512 NC_019725	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. sucinogenes A. hydrophila A. salmonicida A. macleodii B. aphidicola Cand. Blochmannia Cand. Blochmannia Cand. Moranella C. koseri Cand Moranella 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_014977 NC_015172 NC_014220 NC_014220 NC_014220 NC_014220 NC_014202 NC_014964 NC_013921 NC_014964 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_014519 NC_017992 NC_014152 NC_010320 NC_014538 NC_003869 NC_014555 NC_014555 NC_015555 NC_013740 NC_015873 NC_017068 NC_017069 NC_017073	 n. modestication H. orenii H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. motockii T. italicus T. marianensis T. motockii T. italicus T. marianensis T. pocens T. oceani T. phaeum T. phaeum T. potens T. poseudethanolicus Thermoanaerobacter sp. T. tengcongensis T. thermonaerobacter sp. T. tengcongensis T. thermosaccharolyticum A. intestini M. elsdenii S. ruminantium S. ruminantium S. ruminantium 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009555 NC_009348 NC_011312 NC_018690 NC_011312 NC_018679 NC_018673 NC_018673 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018513 NC_004545 NC_004545 NC_004545 NC_004545 NC_004545 NC_0017252 NC_017252 NC_017253 NC_017255 NC_017255 NC_017256 NC_017256 NC_017256 NC_017256 NC_017256 NC_017256 NC_017256 NC_017257 NC_007984 NC_007921 NC_012751 NC_009792 NC_012751 NC_009792 NC_012753	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. salmonicida A. suis A. shrophilus A. fischeri A. macleodii B. aphidicola Cand. Blochmannia Cand. Hamiltonella C. severi C. anverberguther ce 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_0104833 NC_01954 NC_014833 NC_014200 NC_014200 NC_014200 NC_014200 NC_014921 NC_014991 NC_014377 NC_015499 NC_014377 NC_015499 NC_014152 NC_014321 NC_014321 NC_014338 NC_014555 NC_013555 NC_013740 NC_017068 NC_017069 NC_017073 NC_013520	 n. modestication H. orenii H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. mathranii T. narugense T. oceani T. phaeum T. phaeum T. phaeum T. piaeum T. potens T. pseudethanolicus Thermoanaerobacter sp. Thermoanaerobacter sp. T. thermosacharolyticum T. wiegelii T. xylanolyticum A. intestini M. elsdenii S. ruminantium S. sputigena V. parvule 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00939 NC_00955 NC_009348 NC_011312 NC_018690 NC_012913 NC_016853 NC_011318 NC_016853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_015554 NC_002528 NC_004513 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017256 NC_017255 NC_017256 NC_017259 NC_017256 NC_017259 NC_017257 NC_007984 NC_005061 NC_007929 NC_012751 NC_009792 NC_013755 NC_003910	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. suis A. ephrophilus A. fischeri A. macleodii B. aphidicola Cand. Blochmannia Cand. Blochmannia Cand. Hamiltonella C. koseri Cand. Moranella C. psychrerythraea Cand. Biesin 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_014833 NC_015977 NC_014833 NC_015977 NC_014220 NC_014200 NC_016040 NC_008346 NC_013921 NC_01499 NC_01499 NC_014377 NC_018870 NC_014317 NC_014312 NC_014321 NC_014321 NC_014328 NC_014525 NC_013555 NC_013740 NC_017069 NC_017073 NC_015437 NC_013525	 n. modestication H. orenii H. praevalens M. australiensis M. thermophilus N. thermophilus O. valericigenes O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. matraensis T. matraensis T. phage T. potens T. potens T. peseudethanolicus Thermoanaerobacter sp. T. thermoanaerobacter sp. T. tengcongensis T. theremtans A. intestini M. elsdenii S. ruminantium S. sputigena V. parvula T. terrenum 	Clostridia Clostridia	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_00955 NC_00955 NC_009348 NC_011312 NC_018690 NC_015424 NC_012913 NC_016853 NC_011318 NC_016853 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_002528 NC_004061 NC_002528 NC_004545 NC_002528 NC_007253 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_007984 NC_005061 NC_007921 NC_012751 NC_003910 NC_013716	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. suis A. veronii A. aphrophilus A. fischeri A. macleodii B. aphidicola Cand. Blochmannia Cand. Hamiltonella C. koseri Cand. Moranella C. psychrerythraea Cand. Riesia C. rodertium 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010718 NC_010718 NC_010718 NC_010646 NC_016046 NC_01646 NC_015172 NC_015172 NC_015172 NC_015172 NC_015172 NC_015172 NC_015173 NC_015179 NC_015179 NC_014964 NC_015519 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_014377 NC_018870 NC_017992 NC_014538 NC_010320 NC_014538 NC_015555 NC_015555 NC_015873 NC_015637 NC_015437 NC_015437 NC_015437 NC_015437 NC_013526	 n. modessicalization H. orenti H. orenti H. praevalens M. australiensis M. thermopatilus N. thermophilus N. thermophilus O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. marianensis T. marianensis T. marianensis T. potens T. potens T. pesudethanolicus Thermoanaerobacter sp. T. tengcongensis T. thremosacharolyticum A. fermentans A. intestini S. ruminantium S. ruminantium S. ruminantium S. sputigena V. parvula T. terrenum 	Clostridia Negativicutes Negativicutes Negativicutes Negativicutes Chloroflexi Chloroflexi	$\begin{array}{c} 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009359 NC_009348 NC_011312 NC_015424 NC_011312 NC_015424 NC_012913 NC_016853 NC_012913 NC_018678 NC_018678 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_01933 NC_01554 NC_004545 NC_004545 NC_004545 NC_004545 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_007984 NC_005061 NC_00971 NC_003710 NC_013716 NC_003710	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. macleodii B. aphidicola Cand. Blochmannia Cand. Blochmannia Cand. Blochmannia Cand. Moranella C. psychrerythraea Cand. Riesia C. rodentium C. sakazakii 	Pproteobacteria	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010718 NC_010724 NC_016046 NC_016048 NC_009454 NC_014833 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_01420 NC_01494 NC_013519 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_01499 NC_014538 NC_003869 NC_014538 NC_003869 NC_014538 NC_015555 NC_013740 NC_015955 NC_013768 NC_017068 NC_017069 NC_017437 NC_015437 NC_015437 NC_013520 NC_013526 NC_013526	 n. modesticularian H. orenii H. orenii H. praevalens M. australiensis M. thermoacetica N. thermophilus N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. motockii T. norensis T. motockii T. italicus T. marianensis T. pocenii T. phaeum T. phaeum T. potens T. pseudethanolicus Thermoanaerobacter sp. Thermosaccharolyticum T. wiegelii T. zylanolyticum A. fermentans A. intestini M. elsdenii S. ruminantium S. ruminantium S. sputigena V. parvula T. terrenum T. terrenum T. terrenum 	Clostridia Clostridia	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 1.00\\$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_010278 NC_010939 NC_009555 NC_009348 NC_011312 NC_018650 NC_001312 NC_018670 NC_012913 NC_012913 NC_012913 NC_012913 NC_0128673 NC_018673 NC_018679 NC_018679 NC_018679 NC_018679 NC_018554 NC_002528 NC_004545 NC_004545 NC_004545 NC_004545 NC_004545 NC_004545 NC_0017252 NC_017252 NC_017252 NC_017255 NC_017251 NC_009718 NC_003910 NC_014109 NC_013716 NC_009778	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. succinogenes A. salmonicida A. salmonicida A. salmonicida A. salmonicida A. suis A. suis A. ophrophilus A. fischeri A. macleodii B. aphidicola Cand. Blochmannia Cand. Blochmannia Cand. Hamiltonella C. koseri Cand. Moranella C. rodentium C. sakazakii C. sakazakii 	Pproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_011899 NC_011899 NC_017455 NC_015520 NC_007644 NC_010718 NC_010714 NC_010724 NC_016046 NC_016048 NC_019454 NC_014833 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014220 NC_014921 NC_014921 NC_01499 NC_014377 NC_015499 NC_014152 NC_014152 NC_010321 NC_014555 NC_013510 NC_017068 NC_017068 NC_017069 NC_017073 NC_017069 NC_017073 NC_017069 NC_0170520 NC_013526 NC_014836 NC_014836	 n. modesticularin H. orenii H. orenii H. praevalens M. australiensis M. thermopatius N. thermophilus N. thermophilus O. valericigenes O. valericigenes P. thermopropionicum R. albus R. hominis S. glycolicus S. lipocalidus S. thermotolerans S. wolfei T. acetatoxydans T. brockii T. italicus T. marianensis T. motockii T. narugense T. oceani T. phaeum T. phaeum T. phaeum T. potens T. pseudethanolicus Thermoanaerobacter sp. Thermosaccharolyticum T. wiegelii T. xylanolyticum A. intestini M. elsdenii S. ruminantium S. ruminantium S. sputigena V. parvula T. terrenum T. terrenum Synechococcus sp. D. indicum 	Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.86 0.62 1.00	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009053 NC_009053 NC_00939 NC_00955 NC_00939 NC_00955 NC_009348 NC_011312 NC_018670 NC_012913 NC_016853 NC_011312 NC_016853 NC_011832 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_018679 NC_015554 NC_002528 NC_004513 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017255 NC_017256 NC_017259 NC_017255 NC_017256 NC_017259 NC_017255 NC_017257 NC_017257 NC_017257 NC_007984 NC_005061 NC_007929 NC_014099 NC_014109 NC_013716 NC_009778 NC_0133282	 A. pleuropneumoniae A. pleuropneumoniae A. pleuropneumoniae A. succinogenes A. submonicida A. salmonicida A. salmonicida A. salmonicida A. suis A. suis A. ophrophilus A. fischeri A. macleodii B. aphidicola Cand. Blochmannia Cand. Blochmannia Cand. Horhannia Cand. Moranella C. rodentium C. sakazakii C. sakazakii C. sakazakii C. sukazakii C. sukazakii C. sukazakii C. sukazakii C. sukazakii 	Pproteobacteria	$\begin{array}{c} 1.00\\$

CHR	NC_{012880}	D. dadantii	γ proteobacteria	1.00	PL	NC_011409	H. influenzae	γ proteobacteria	1.00
CHR	NC_013592	D. dadantii	γ proteobacteria	1.00	CHR	NC_014920	H. influenzae	γ proteobacteria	1.00
CHR	NC_014500	D. dadantii	γ proteobacteria	1.00	CHR	NC_014922	H. influenzae	γ proteobacteria	1.00
CHR	NC 012912	D. nouosus D. zeae	vproteobacteria	1.00	CHR	NC 017451	H influenzae	γproteobacteria	1.00
CHR	NC_015663	E. aerogenes	γ proteobacteria	1.00	CHR	NC_017452	H. influenzae	γ proteobacteria	1.00
PL	NC_005247	E. amylovora	γ proteobacteria	0.73	CHR	NC_015964	H. parainfluenzae	γ proteobacteria	1.00
CHR	NC_013961	E. amylovora	γ proteobacteria	1.00	CHR	NC_011852	H. parasuis	γ proteobacteria	1.00
CHR	NC_013971	E. amylovora	γ proteobacteria	1.00	CHR	NC_008309	H. somnus	γ proteobacteria	1.00
PI	NC_015968	E. asouriae E. asburiae	γproteobacteria	1.00	CHR	NC 006512	H. somnus L. loibiensis	γ proteobacteria	1.00
CHR	NC_014306	E. billingiae	γproteobacteria	nan	CHR	NC_013166	K. koreensis	γproteobacteria	1.00
CHR	NC_017910	E. blattae	γ proteobacteria	1.00	CHR	NC_016612	K. oxytoca	γ proteobacteria	1.00
CHR	NC_{014121}	E. cloacae	γ proteobacteria	1.00	CHR	NC_018106	K. oxytoca	γ proteobacteria	1.00
CHR	NC_014618	E. cloacae	γ proteobacteria	1.00	CHR	NC_009648	K. pneumoniae	γ proteobacteria	1.00
CHR	NC_016514	E. cloacae E. cloacae	γ proteobacteria	1.00	CHR	NC_011283 NC_012731	K. pneumoniae K. pneumoniae	γproteobacteria	1.00
CHR	NC_018405	E. cloacae	vproteobacteria	1.00	CHR	NC_012731 NC_016845	K. pneumoniae	γproteobacteria	1.00
CHR	NC_000913	E. coli	γ proteobacteria	1.00	CHR	NC_017540	K. pneumoniae	γ proteobacteria	1.00
CHR	$NC_{-}002655$	E. coli	γ proteobacteria	1.00	CHR	NC_018522	K. pneumoniae	γ proteobacteria	1.00
CHR	NC_002695	E. coli	γ proteobacteria	1.00	CHR	NC_013850	K. variicola	γ proteobacteria	1.00
CHR	NC_004431	E. coli E. coli	γ proteobacteria	1.00	CHR	NC_014147	M. catarrhalis	γproteobacteria	1.00
CHR	NC_007946	E. coli	vproteobacteria	1.00	CHR	NC_017856	M. guvum Methylophaga sp.	vproteobacteria	0.12
CHR	NC_008253	E. coli	γ proteobacteria	1.00	CHR	NC_006300	M. succiniciproducens	γ proteobacteria	1.00
CHR	NC_008563	E. coli	γ proteobacteria	1.00	CHR	$NC_{-}016745$	Oceanimonas sp.	γ proteobacteria	1.00
CHR	NC_009800	E. coli	γ proteobacteria	1.00	CHR	NC_002663	P. multocida	γ proteobacteria	1.00
CHR	NC_009801	E. coli E. coli	γ proteobacteria	1.00	CHR	NC_016808	P. multocida P. multocida	γ proteobacteria	nan
CHR	NC_010408	E. coli	vproteobacteria	1.00	CHR	NC_017027	P. multocida	vproteobacteria	1.00
CHR	NC_010498	E. coli	γ proteobacteria	1.00	CHR	NC_013956	P. ananatis	γ proteobacteria	1.00
CHR	NC_{011353}	E. coli	γ proteobacteria	1.00	CHR	NC_016816	P. ananatis	γ proteobacteria	1.00
CHR	NC_011415	E. coli	γ proteobacteria	1.00	CHR	NC_017531	P. ananatis	γ proteobacteria	1.00
CHR	NC_011601	E. coli	γ proteobacteria	1.00	CHR	NC_017554	P. ananatis	γ proteobacteria	1.00
CHR	NC 011741 NC 011742	E. coli E. coli	γproteobacteria	1.00	CHR	NC 004547	P. atrosenticum	γproteobacteria	1.00
CHR	NC_011745	E. coli	γ proteobacteria	1.00	CHR	NC_012917	P. carotovorum	γ proteobacteria	1.00
CHR	NC_011748	E. coli	γ proteobacteria	1.00	CHR	NC_{018525}	P. carotovorum	γ proteobacteria	1.00
CHR	NC_011750	E. coli	γ proteobacteria	1.00	CHR	NC_006370	P. profundum	γ proteobacteria	1.00
CHR	NC_011751	E. coli	γ proteobacteria	1.00	CHR	NC_012962	P. asymbiotica	γ proteobacteria	1.00
CHR	NC 012759	E. coli E. coli	γproteobacteria	1.00	CHR	NC 0105126	P. iuminescens P. mirahilis	γproteobacteria	1.00
CHR	NC_012892	E. coli	γproteobacteria	1.00	CHR	NC_008228	P. atlantica	γ proteobacteria	1.00
CHR	NC_012947	E. coli	γ proteobacteria	1.00	CHR	NC_007969	P. cryohalolentis	γ proteobacteria	1.00
CHR	NC_012967	E. coli	γ proteobacteria	1.00	CHR	NC_007481	P. haloplanktis	γ proteobacteria	1.00
CHR	NC_012971	E. coli	γ proteobacteria	1.00	RECE	NC_007482	P. haloplanktis	γ proteobacteria	0.95
				1 1 00	DECE	NCL 014800	Described and the second second second		
CHR	NC_013008 NC_013353	E. coli E. coli	γ proteobacteria γ proteobacteria	1.00	RECE CHB	NC_014800 NC_014803	Pseudoalteromonas sp. Pseudoalteromonas sp	γ proteobacteria γ proteobacteria	1.00
CHR CHR CHR	NC_013008 NC_013353 NC_013361	E. coli E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00	RECE CHR CHR	NC_014800 NC_014803 NC_007204	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus	γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00
CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364	E. coli E. coli E. coli E. coli	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00 1.00	RECE CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp.	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00 1.00
CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654	E. coli E. coli E. coli E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 $	RECE CHR CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524 NC_008709	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00 1.00
CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013941	E. coli E. coli E. coli E. coli E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 1.00$	RECE CHR CHR CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524 NC_008709 NC_017731	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00 1.00 1.00
CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013941 NC_016902 NC_017625	E. coli E. coli E. coli E. coli E. coli E. coli E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 1.00$	RECE CHR CHR CHR CHR CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524 NC_008709 NC_017731 NC_014562 NC_013421	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. stuartii P. vagans P. wagans	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013941 NC_016902 NC_017625 NC_017625	E. coli E. coli E. coli E. coli E. coli E. coli E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00 \end{array} $	RECE CHR CHR CHR CHR CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524 NC_008709 NC_017731 NC_014562 NC_013421 NC_016818	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. stuartii P. vagans P. wasabiae R. aquatilis	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00 \end{array} $
CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013654 NC_013654 NC_013941 NC_016902 NC_017625 NC_017626 NC_017628	E. coli E. coli E. coli E. coli E. coli E. coli E. coli E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	RECE CHR CHR CHR CHR CHR CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524 NC_008709 NC_017731 NC_014562 NC_013421 NC_016818 NC_017047	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00 \end{array} $
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013941 NC_016902 NC_017625 NC_017626 NC_017628 NC_017631	E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524 NC_008709 NC_017731 NC_014562 NC_013421 NC_016818 NC_017047 NC_015061	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp.	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_016902 NC_017625 NC_017626 NC_017631 NC_017631 NC_017632	E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014800 NC_014803 NC_007204 NC_009524 NC_008709 NC_017731 NC_014562 NC_013421 NC_016818 NC_017047 NC_015061 NC_008700	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. kelvis	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_016902 NC_016902 NC_017625 NC_017626 NC_017631 NC_017633 NC_017634	E. coli E. coli	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.007204 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.015061 NC.008700 NC.009665 NC.009997	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. baltica	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013005 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_016902 NC_017625 NC_017625 NC_017628 NC_017631 NC_017632 NC_017633	E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.011663	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis R. aquatilis R. aquatilis S. banzonensis S. baltica S. baltica	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013005 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_016902 NC_017625 NC_017626 NC_017631 NC_017632 NC_017633 NC_017634 NC_017635 NC_017638	E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.011663 NC.016901	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. baltica S. baltica	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013005 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_016902 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017634 NC_017635 NC_017641	E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.011663 NC.016901 NC.017571	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. baltica S. baltica S. baltica	γproteobacteria	$\begin{array}{c} 1.00\\$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013005 NC_013353 NC_013361 NC_013364 NC_013364 NC_013654 NC_013941 NC_016902 NC_017625 NC_017626 NC_017631 NC_017633 NC_017634 NC_017634 NC_017641 NC_017644	E. coli E. coli	γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria γρroteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.011663 NC.016901 NC.017571 NC.015761	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. baltica	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013005 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_013654 NC_017625 NC_017625 NC_017631 NC_017633 NC_017633 NC_017633 NC_017634 NC_017641 NC_017644 NC_017646 NC_017651	E. coli E. coli	γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.01663 NC.016901 NC.017571 NC.015761 NC.01658	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. baltica	γproteobacteria	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013005 NC_013353 NC_013361 NC_013364 NC_013364 NC_013654 NC_013654 NC_017625 NC_017625 NC_017631 NC_017633 NC_017633 NC_017633 NC_017635 NC_017641 NC_017644 NC_017651 NC_017655	E. coli E. coli	γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.008700 NC.009665 NC.009997 NC.01665 NC.016901 NC.017571 NC.015761 NC.007613 NC.007654	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. baltica	γproteobacteria	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013361 NC_013654 NC_013941 NC_016902 NC_017625 NC_017626 NC_017631 NC_017633 NC_017633 NC_017634 NC_017635 NC_017644 NC_017644 NC_017646 NC_017655 NC_017655 NC_017656	E. coli E. coli	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.01663 NC.017571 NC.017571 NC.017571 NC.017571 NC.007613 NC.007954 NC.007606	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. bolydii S. boydii S. denitrificans S. dysenteriae	γproteobacteria	1.00 1.000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013363 NC_013361 NC_013364 NC_013654 NC_013941 NC_016902 NC_017625 NC_017626 NC_017631 NC_017633 NC_017633 NC_017633 NC_017634 NC_017635 NC_017644 NC_017644 NC_017656 NC_017656 NC_017656 NC_017656	E. $coli$ E. $coli$	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014800 NC_007204 NC_009524 NC_009524 NC_009524 NC_014562 NC_013421 NC_016818 NC_017047 NC_015061 NC_009665 NC_009665 NC_009697 NC_01663 NC_017571 NC_017571 NC_017571 NC_017571 NC_007613 NC_007634 NC_007606 NC_007606	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. baltica	Proteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013364 NC_013654 NC_013941 NC_016902 NC_017625 NC_017626 NC_017631 NC_017633 NC_017634 NC_017635 NC_017634 NC_017644 NC_017644 NC_017651 NC_017656 NC_017656 NC_017656 NC_017660 NC_017660 NC_017660	E. coli E.	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009907 NC.011663 NC.016901 NC.017571 NC.01658 NC.007613 NC.007613 NC.010658 NC.007954 NC.007606 NC.018419 NC.018420	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. boldii S. boydii S. doydii S. doydii S. dysenteriae s. endosymbiont S. endosymbiont	Pproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_016902 NC_016902 NC_017625 NC_017626 NC_017631 NC_017633 NC_017633 NC_017633 NC_017634 NC_017644 NC_017655 NC_017655 NC_017656 NC_017666 NC_017666 NC_017663 NC_017664 NC_017664 NC_017664	E. coli	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009907 NC.011663 NC.016901 NC.017571 NC.015613 NC.007613 NC.007613 NC.007658 NC.007954 NC.007606 NC.018419 NC.018419 NC.003197	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. bolgdii S. doydii S. doydii S. dysenteriae s. endosymbiont S. enterica	Pproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013053 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_017625 NC_017625 NC_017626 NC_017632 NC_017631 NC_017633 NC_017633 NC_017634 NC_017638 NC_017638 NC_017651 NC_017651 NC_017651 NC_017656 NC_017666 NC_017663 NC_017663 NC_017664 NC_017664 NC_017665 NC_017663	E. $coli$ E. $coli$	γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009605 NC.009907 NC.011663 NC.01663 NC.01663 NC.007613 NC.007613 NC.007613 NC.007658 NC.007954 NC.007606 NC.018419 NC.018419 NC.003197 NC.003197 NC.003198 NC.004631	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. balti	yproteobacteria	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013053 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_0136902 NC_017625 NC_017626 NC_017632 NC_017631 NC_017633 NC_017633 NC_017633 NC_017638 NC_017653 NC_017651 NC_017651 NC_017651 NC_017656 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017665 NC_018655	E. $coli$ E. $coli$	proteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009605 NC.009997 NC.011663 NC.016901 NC.017571 NC.01663 NC.007613 NC.007613 NC.007606 NC.018420 NC.003197 NC.003197 NC.003197 NC.003191	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. centerica S. enterica	proteobacteria proteobacteria proteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013005 NC_013353 NC_013361 NC_013364 NC_013364 NC_013654 NC_013654 NC_016902 NC_017625 NC_017626 NC_017632 NC_017633 NC_017633 NC_017633 NC_017635 NC_017638 NC_017651 NC_017651 NC_017651 NC_017656 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017665 NC_018655 NC_018655 NC_018651	E. $coli$ E. $coli$	proteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009605 NC.009997 NC.011663 NC.016901 NC.017571 NC.01651 NC.007613 NC.007613 NC.010658 NC.007954 NC.007606 NC.018420 NC.003197 NC.003197 NC.003197 NC.003191 NC.004631 NC.006511 NC.006905	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. canterica S. enterica S. enterica	proteobacteria proteobacteria pr	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013035 NC_013353 NC_013361 NC_013361 NC_013364 NC_013654 NC_017625 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017633 NC_017634 NC_017635 NC_017644 NC_017644 NC_017646 NC_017655 NC_017655 NC_017655 NC_017656 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_018656 NC_018658 NC_018658	E. $coli$ E. $coli$	proteobacteria	$\begin{array}{c} 1.00\\ 0.00\\ 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.01663 NC.016901 NC.017571 NC.01658 NC.007613 NC.010658 NC.007606 NC.018420 NC.003197 NC.003197 NC.003198 NC.004631 NC.006905 NC.010067	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. chterica S. enterica S. enterica S. enterica	γproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013361 NC_013361 NC_013361 NC_013364 NC_017625 NC_017625 NC_017631 NC_017631 NC_017633 NC_017633 NC_017634 NC_017635 NC_017641 NC_017651 NC_017652 NC_017656 NC_017656 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017664 NC_017665 NC_017664 NC_017665 NC_018656 NC_018656 NC_018651 NC_012779 NC_012214	E. $coli$ E. $coli$	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.016818 NC.017047 NC.015061 NC.008700 NC.009665 NC.009997 NC.01663 NC.016901 NC.017571 NC.015761 NC.01658 NC.007606 NC.010658 NC.007954 NC.007606 NC.018420 NC.003197 NC.003197 NC.003198 NC.006905 NC.010067 NC.010067 NC.011080	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. cherica S. enterica S. enterica S. enterica S. enterica	Proteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013363 NC_013364 NC_013364 NC_013654 NC_013644 NC_016902 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017633 NC_017633 NC_017634 NC_017635 NC_017638 NC_017644 NC_017644 NC_017646 NC_017656 NC_017656 NC_017666 NC_017666 NC_017666 NC_017664 NC_017664 NC_017664 NC_017664 NC_017664 NC_017665 NC_017664 NC_017665 NC_017664 NC_017664 NC_017665 NC_018658 NC_018658 NC_018658 NC_018679 NC_012779 NC_012214	E. $coli$ E. $coli$	proteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.008709 NC.017731 NC.014562 NC.013421 NC.014562 NC.013421 NC.016818 NC.015061 NC.009097 NC.015661 NC.009997 NC.015761 NC.015761 NC.015761 NC.015761 NC.007605 NC.007606 NC.018419 NC.003197 NC.003198 NC.004631 NC.006511 NC.006905 NC.010067 NC.01002 NC.011080	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. cherica S. enterica S. enterica S. enterica S. enterica S. enterica	Pproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_016902 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017633 NC_017634 NC_017635 NC_017635 NC_017644 NC_017644 NC_017646 NC_017656 NC_017656 NC_017660 NC_017660 NC_017663 NC_017663 NC_017663 NC_017664 NC_017664 NC_017664 NC_017664 NC_017665 NC_018650 NC_018650 NC_018650 NC_018651 NC_011740 NC_012779 NC_002214 NC_017390 NC_009436	E. $coli$ E. $coli$	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.017204 NC.007204 NC.007204 NC.007204 NC.017731 NC.014562 NC.014562 NC.014562 NC.016818 NC.017047 NC.015061 NC.009907 NC.01663 NC.009907 NC.011663 NC.016581 NC.007613 NC.010658 NC.007954 NC.003197 NC.018419 NC.018420 NC.003197 NC.003198 NC.006511 NC.006905 NC.010067 NC.01002 NC.011080 NC.011083 NC.011094	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. cherica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica	Pproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_013654 NC_016902 NC_017625 NC_017625 NC_017631 NC_017631 NC_017632 NC_017633 NC_017634 NC_017635 NC_017635 NC_017644 NC_017644 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017665 NC_017665 NC_017665 NC_017664 NC_017665 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018651 NC_011740 NC_012779 NC_009436 NC_009436 NC_017445	E. $coli$ E. $coli$	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.017204 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.014562 NC.017047 NC.015061 NC.008700 NC.009907 NC.015661 NC.009907 NC.011663 NC.01663 NC.016630 NC.016518 NC.007614 NC.007614 NC.007616 NC.001954 NC.001954 NC.001954 NC.004631 NC.006511 NC.006905 NC.010067 NC.01007 NC.010102 NC.011083 NC.011094 NC.011047	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. chererica S. enterica S. enterica	yproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013364 NC_013654 NC_013654 NC_017625 NC_017625 NC_017626 NC_017631 NC_017632 NC_017633 NC_017633 NC_017634 NC_017635 NC_017634 NC_017635 NC_017644 NC_017651 NC_017656 NC_017660 NC_017660 NC_017660 NC_017663 NC_017660 NC_017663 NC_017664 NC_017660 NC_017663 NC_017664 NC_017666 NC_017664 NC_017666 NC_017666 NC_017664 NC_017666 NC_017664 NC_017664 NC_017660 NC_018650 NC_018650 NC_018650 NC_018650 NC_018651 NC_012779 NC_012214 NC_012790 NC_009436 NC_017445 NC_013508	E. coli E. fergusonii E. fergusonii E. pyrifoliae E. pyrifoliae E. pyrifoliae E. tarda	Proteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.007204 NC.009524 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.01661 NC.008700 NC.009065 NC.00997 NC.015061 NC.007603 NC.01663 NC.01663 NC.01663 NC.007613 NC.016658 NC.007613 NC.007613 NC.007613 NC.007613 NC.007613 NC.007613 NC.007606 NC.018419 NC.003197 NC.003197 NC.003197 NC.004631 NC.004631 NC.006511 NC.006905 NC.010002 NC.011083 NC.011094 NC.011147 NC.011147	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. chterica S. enterica S. enterica	proteobacteria proteobacteria pr	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013053 NC_013353 NC_013361 NC_013364 NC_013364 NC_013654 NC_013654 NC_017625 NC_017625 NC_017626 NC_017632 NC_017633 NC_017633 NC_017633 NC_017638 NC_017638 NC_017638 NC_017641 NC_017651 NC_017650 NC_017650 NC_017660 NC_017663 NC_017663 NC_017663 NC_017663 NC_017665 NC_017665 NC_017665 NC_017665 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_012779 NC_012214 NC_017390 NC_009436 NC_017309	E. $coli$ E. $coli$	Proteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.015061 NC.008700 NC.009097 NC.01561 NC.009097 NC.011663 NC.01663 NC.01663 NC.01663 NC.007613 NC.016658 NC.007613 NC.007613 NC.007613 NC.00761420 NC.018419 NC.018419 NC.003197 NC.003198 NC.004631 NC.006511 NC.006511 NC.006511 NC.006511 NC.006511 NC.006051 NC.01002 NC.011080 NC.011080 NC.011080 NC.011047	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. enterica S. enterica	Proteobacteria	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013053 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_017625 NC_017625 NC_017626 NC_017632 NC_017631 NC_017633 NC_017633 NC_017633 NC_017638 NC_017638 NC_017638 NC_017641 NC_017651 NC_017651 NC_017655 NC_017666 NC_017666 NC_017665 NC_017666 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_018650 NC_018650 NC_018658 NC_018650 NC_018650 NC_018650 NC_018650 NC_012779 NC_012214 NC_017390 NC_009436 NC_017445 NC_017309 NC_017309	E. $coli$ E. $coli$	Proteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.016818 NC.016818 NC.016070 NC.009097 NC.011663 NC.016901 NC.017571 NC.01663 NC.01663 NC.007613 NC.007613 NC.007613 NC.007613 NC.00761420 NC.018419 NC.018420 NC.003197 NC.0018420 NC.003198 NC.004631 NC.006511 NC.006511 NC.006511 NC.006511 NC.006051 NC.01002 NC.011080 NC.011080 NC.011080 NC.011081 NC.011147 NC.011294	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis R. aquatilis R. aquatilis R. aquatilis R. aquatilis R. aquatilis S. baltica S. enterica S. enterica	proteobacteria proteobacteria pr	1.00 1.00
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013035 NC_013353 NC_013361 NC_013361 NC_013364 NC_013654 NC_017625 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017633 NC_017634 NC_017635 NC_017634 NC_017644 NC_017644 NC_017646 NC_017651 NC_017655 NC_017655 NC_017656 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017790 NC_012214 NC_017390 NC_009436 NC_017309 NC_017309 NC_0105460	E. $coli$ E. $coli$	proteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.01661 NC.008700 NC.009605 NC.009907 NC.011663 NC.016901 NC.017571 NC.01663 NC.007613 NC.007613 NC.007613 NC.007651 NC.007606 NC.018420 NC.003197 NC.003197 NC.003198 NC.004631 NC.006511 NC.006511 NC.006511 NC.006305 NC.01002 NC.011080 NC.011083 NC.011083 NC.011094 NC.01125	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. enterica S. enterica	γproteobacteria γproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013361 NC_013364 NC_013654 NC_017625 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017633 NC_017634 NC_017635 NC_017644 NC_017644 NC_017644 NC_017652 NC_017652 NC_017656 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017667 NC_018650 NC_018658 NC_018661 NC_01740 NC_012214 NC_017309 NC_009436 NC_017309 NC_016941 NC_015460 NC_016041	E. $coli$ E. $coli$	yproteobacteria yproteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.013421 NC.015061 NC.008700 NC.009665 NC.009907 NC.01663 NC.016901 NC.017571 NC.01663 NC.016901 NC.017571 NC.01658 NC.007613 NC.007613 NC.010658 NC.007954 NC.007606 NC.0108420 NC.003197 NC.003197 NC.003197 NC.003197 NC.003197 NC.003197 NC.004631 NC.006511 NC.006511 NC.006511 NC.006905 NC.01002 NC.011080 NC.011083 NC.011083 NC.011094 NC.01125 NC.011275 NC.011275 NC.011275	$\begin{tabular}{lllllllllllllllllllllllllllllllllll$	yproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013363 NC_013364 NC_013364 NC_013654 NC_013644 NC_016902 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017633 NC_017633 NC_017634 NC_017635 NC_017635 NC_017636 NC_017664 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017667 NC_012779 NC_012214 NC_017390 NC_009436 NC_017455 NC_017309 NC_00694 NC_016694 NC_016041 NC_01641 NC_01641	E. $coli$ E. $coli$	Proteobacteria	$\begin{array}{c} 1.00\\ 0.00\\ 1.00\\ 0.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.014562 NC.015061 NC.008700 NC.009097 NC.015661 NC.009997 NC.01663 NC.01663 NC.016761 NC.007613 NC.016581 NC.007606 NC.015761 NC.007606 NC.015761 NC.007606 NC.015761 NC.007606 NC.015761 NC.007606 NC.018419 NC.003198 NC.004631 NC.006905 NC.010067 NC.010083 NC.011094 NC.011094 NC.011147 NC.011205 NC.011274 NC.011274 NC.01225 NC.016810 NC.016831	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. chterica S. enterica S. enterica	yproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_013654 NC_017625 NC_017626 NC_017626 NC_017631 NC_017631 NC_017633 NC_017633 NC_017634 NC_017635 NC_017635 NC_017635 NC_017664 NC_017664 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017665 NC_017665 NC_017665 NC_017665 NC_017666 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017665 NC_017664 NC_017665 NC_017665 NC_017665 NC_017664 NC_017300 NC_009436 NC_017455 NC_017309 NC_00694 NC_016041 NC_015497 NC_002940	E. $coli$ E. $coli$	Proteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.008709 NC.017731 NC.014562 NC.013421 NC.014562 NC.016818 NC.017047 NC.015061 NC.009907 NC.01663 NC.009907 NC.011663 NC.016630 NC.016581 NC.007606 NC.018419 NC.016581 NC.006905 NC.010067 NC.010102 NC.010087 NC.010083 NC.011205 NC.011205 NC.011274 NC.011205 NC.011274 NC.011215 NC.016831 NC.016831 NC.016832 NC.016831 NC.016832	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. boydii S. doydii S. doydii S. doydii S. doydii S. doydii S. doydii S. doydii S. denitrificans S. dysenteriae S. enterica S. en	yproteobacteria	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_017625 NC_017625 NC_017626 NC_017631 NC_017631 NC_017633 NC_017634 NC_017635 NC_017634 NC_017635 NC_017635 NC_017664 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017666 NC_017665 NC_017665 NC_017665 NC_017665 NC_017666 NC_017665 NC_017666 NC_017666 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_01740 NC_012779 NC_009436 NC_017445 NC_013508 NC_017445 NC_013641 NC_015497 NC_002940 NC_000907	E. $coli$ E. $coli$	Proteobacteria	$\begin{array}{c} 1.00\\$	RECCE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.014803 NC.007204 NC.009524 NC.009524 NC.017731 NC.014562 NC.014562 NC.013421 NC.014562 NC.017047 NC.015061 NC.009907 NC.015061 NC.009907 NC.01663 NC.01663 NC.007613 NC.016658 NC.007614 NC.007616 NC.018419 NC.018419 NC.018420 NC.016851 NC.011294 NC.01102 NC.01102 NC.01102 NC.01104 NC.01104 NC.011294 NC.011274 NC.011274 NC.01225 NC.0168510 NC.016854 NC.016854	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis Rahnella sp. S. amazonensis S. baltica S. chererica S. enterica S. enterica S	Pproteobacteriapproteobacteria </td <td>1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1</td>	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_013008 NC_013353 NC_013361 NC_013364 NC_013654 NC_013654 NC_013654 NC_017625 NC_017625 NC_017628 NC_017631 NC_017631 NC_017632 NC_017633 NC_017634 NC_017635 NC_017635 NC_017644 NC_017644 NC_017646 NC_017660 NC_017660 NC_017660 NC_017660 NC_017663 NC_017660 NC_017663 NC_017664 NC_017664 NC_017664 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_018650 NC_01740 NC_012779 NC_009436 NC_017445 NC_013508 NC_016941 NC_015497 NC_009466	E. $coli$ E. $coli$	Proteobacteria	$\begin{array}{c} 1.00\\$	RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014800 NC.017204 NC.009524 NC.009524 NC.007204 NC.017731 NC.014562 NC.013421 NC.014562 NC.017047 NC.015061 NC.009907 NC.015061 NC.009907 NC.011663 NC.01663 NC.01663 NC.007613 NC.007613 NC.007613 NC.007613 NC.007613 NC.007613 NC.007614 NC.007606 NC.018419 NC.004631 NC.004631 NC.006905 NC.010067 NC.01002 NC.01102 NC.01102 NC.01104 NC.01104 NC.011294 NC.011294 NC.016831 NC.016832 NC.016856 NC.016856	Pseudoalteromonas sp. Pseudoalteromonas sp. P. arcticus Psychrobacter sp. P. ingrahamii P. stuartii P. vagans P. wasabiae R. aquatilis R. aquatilis Rahnella sp. S. amazonensis S. baltica S. chererica S. enterica S. enterica	proteobacteria proteobacteria pr	1.000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000 1.0000

Annexe D.Résultats du clustering de V^R par INFOMAP

CHR	NC 016863	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC 017958	T. mobilis	αproteobacteria	0.00
CHR	NC_017046	S. enterica	γ proteobacteria	1.00	PL	NC_017966	T. mobilis	α proteobacteria	0.57
CHR	NC_017623	S. enterica	γ proteobacteria	1.00	CHR	NC_006513	A. aromaticum	β proteobacteria	1.00
CHR	NC_004337	S. flexneri	γ proteobacteria	1.00	CHR	NC_015138	A. avenae	β proteobacteria	1.00
CHR	NC_004741	S. flexneri	γ proteobacteria	1.00	CHR	NC_008752	A. citrulli	β proteobacteria	1.00
CHR	NC_008258	S. flexneri	γproteobacteria	1.00	PL	NC_014908	A. denitrificans	Broteobacteria	0.00
CHR	NC 016632	S. sumbiotica	vproteobacteria	1.00	CHR	NC 015422	A. denitrificans	βproteobacteria	1.00
CHR	NC_015566	Serratia sp.	γproteobacteria	1.00	CHR	NC_011992	A. ebreus	β proteobacteria	1.00
CHR	NC_017573	Serratia sp.	γ proteobacteria	1.00	CHR	NC_017964	A. kashmirensis	β proteobacteria	1.00
CHR	NC_007384	S. sonnei	γ proteobacteria	1.00	CHR	NC_008702	Azoarcus sp.	β proteobacteria	1.00
CHR	NC_016822	S. sonnei	γ proteobacteria	1.00	CHR	NC_008782	Acidovorax sp.	β proteobacteria	1.00
CHR	NC_008345	S. frigidimarina	γ proteobacteria	1.00	CHR	NC_018708	Acidovorax sp.	β proteobacteria	1.00
CHR	NC_007712	S. glossinidius	γ proteobacteria	1.00	CHR	NC_014640	A. xylosoxidans	β proteobacteria	1.00
CHR	NC_010334	S. halifaxensis S. loibica	γproteobacteria	1.00	PL	NC_014642	A. xylosoxidans B. ambifaria	Bproteobacteria	0.47
CHR	NC 004347	S. oneidensis	vproteobacteria	1.00	BECE	NC 008391	B. ambifaria B. ambifaria	Bproteobacteria	0.64
CHR	NC_009901	S. pealeana	γproteobacteria	1.00	RECE*	NC_008392	B. ambifaria	β proteobacteria	0.27
CHR	NC_011566	S. piezotolerans	γ proteobacteria	1.00	\mathbf{CHR}	NC_010551	B. ambifaria	β proteobacteria	1.00
CHR	NC_015567	S. plymuthica	γ proteobacteria	1.00	RECE	NC_010552	B. ambifaria	β proteobacteria	0.47
CHR	NC_009832	S. proteamaculans	γ proteobacteria	1.00	RECE*	NC_010557	B. ambifaria	β proteobacteria	0.47
CHR	NC_009438	S. putrefaciens	γ proteobacteria	1.00	CHR	NC_010645	B. avium	β proteobacteria	1.00
CHR	NC 009831	S. putrejaciens	γproteobacteria	1.00	CHR	NC 018820	B. bronchiseptica	pproteobacteria	1.00
CHR	NC 008321	Shewanella sp.	vproteobacteria	1.00	CHR	NC 019382	B. bronchiseptica	βproteobacteria	1.00
CHR	NC_008322	Shewanella sp.	γ proteobacteria	1.00	CHR	NC_008060	B. cenocepacia	β proteobacteria	1.00
CHR	NC_008577	Shewanella sp.	γ proteobacteria	1.00	RECE	NC_008061	B. cenocepacia	β proteobacteria	0.47
CHR	NC_008750	Shewanella sp.	γ proteobacteria	1.00	RECE*	NC_008062	$B.\ cenocepacia$	β proteobacteria	0.43
CHR	NC_014012	S. violacea	γ proteobacteria	1.00	CHR	NC_008542	B. cenocepacia	β proteobacteria	nan
CHR	NC_010506	S. woodyi	γ proteobacteria	1.00	RECE	NC_008543	B. cenocepacia	β proteobacteria	0.64
CHR	NC_012691	1. auensis T. crunocence	γproteobacteria	0.60	CHP	NC 010509	B. cenocepacia	βproteobacteria	0.00
CHR	NC.015581	T. cyclicum	vproteobacteria	nan	RECE*	NC_010512	B. cenocepacia B. cenocepacia	βproteobacteria	0.00
CHR	NC_015633	V. anguillarum	γproteobacteria	1.00	RECE	NC_010515	B. cenocepacia	β proteobacteria	0.33
CHR	NC_002505	V. cholerae	γ proteobacteria	1.00	CHR	NC_011000	B. cenocepacia	β proteobacteria	nan
CHR	NC_009457	V. cholerae	γ proteobacteria	1.00	RECE	NC_011001	$B.\ cenocepacia$	$\beta \mathrm{proteobacteria}$	0.43
CHR	NC_012578	V. cholerae	γ proteobacteria	1.00	RECE*	NC_011002	B. cenocepacia	β proteobacteria	0.47
CHR	NC_012582	V. cholerae	γ proteobacteria	1.00	CHR	NC_018513	B. cepacia	β proteobacteria	1.00
CHR	NC 016445	v. cholerae V. cholerae	γproteobacteria	1.00	RECE	NC 015376	B. cepacia B. aladioli	pproteobacteria proteobacteria	0.47
CHR	NC_016944	V. cholerae	vproteobacteria	1.00	CHR	NC-015381	B. gladioli B. aladioli	βproteobacteria	1.00
CHR	NC_017270	V. cholerae	γ proteobacteria	1.00	PL	NC_015382	B. gladioli	β proteobacteria	0.00
CHR	NC_006840	V. fischeri	γ proteobacteria	1.00	RECE	NC_012721	B. glumae	β proteobacteria	0.27
CHR	NC_011184	V. fischeri	γ proteobacteria	1.00	CHR	NC_012724	$B. \ glumae$	β proteobacteria	1.00
CHR	NC_016602	V. furnissii	γ proteobacteria	1.00	CHR	NC_006348	B. mallei	β proteobacteria	1.00
CHR	NC_009783	V. harveyi	γproteobacteria	1.00	RECE	NC_006349	B. mallei B. mallei	βproteobacteria	0.47
CHR	NC_013456	Vibrio sp.	vproteobacteria	1.00	CHR	NC_008785	B. mallei	βproteobacteria	nan
CHR	NC_016613	Vibrio sp.	γ proteobacteria	1.00	RECE	NC_008835	B. mallei	β proteobacteria	0.00
CHR	NC_011753	V. splendidus	γ proteobacteria	1.00	\mathbf{CHR}	NC_008836	B. mallei	β proteobacteria	1.00
CHR	NC_004459	V. vulnificus	γ proteobacteria	1.00	RECE	NC_009079	B. mallei	$eta \mathrm{proteobacteria}$	0.47
CHR	NC_005139	V. vulnificus	γ proteobacteria	1.00	CHR	NC_009080	B. mallei	β proteobacteria	1.00
CHR	NC_014965	V. vulnificus	γproteobacteria	1.00	CHR	NC_010084	B. multivorans	βproteobacteria	1.00
CHR	NC 016893	W alossinidia	γproteobacteria	1.00	CHR	NC 010804	B. multivorans	Bproteobacteria	1.00
CHR	NC_013892	X. bovienii	γproteobacteria	1.00	RECE	NC_010805	B. multivorans	βproteobacteria	0.43
CHR	NC_014228	X. nematophila	γ proteobacteria	1.00	CHR	NC_002928	B. parapertussis	β proteobacteria	1.00
CHR	NC_008800	Y. enterocolitica	γ proteobacteria	1.00	CHR	NC_018518	B. parapertussis	β proteobacteria	1.00
CHR	NC_015224	Y. enterocolitica	γ proteobacteria	1.00	CHR	NC_018828	B. parapertussis	β proteobacteria	1.00
CHR	NC_017564	Y. enterocolitica	γ proteobacteria	1.00	CHR	NC_002929	B. pertussis	β proteobacteria	1.00
CHR	NC 004088	Y. pestis V. pestis	γ proteobacteria	1.00	CHR	NC 01017223	B. pertussis B. petrii	pproteobacteria	1.00
CHR	NC_005810	Y. pestis	vproteobacteria	1.00	RECE	NC_018672	B. phenoliruntrix	βproteobacteria	0.00
CHR	NC_008149	Y. pestis	γ proteobacteria	1.00	CHR	NC_018695	B. phenoliruptrix	β proteobacteria	1.00
CHR	NC_008150	Y. pestis	γ proteobacteria	1.00	\mathbf{PL}	NC_018696	B. phenoliruptrix	β proteobacteria	0.00
CHR	NC_009381	Y. pestis	γ proteobacteria	1.00	CHR	NC_010622	B. phymatum	β proteobacteria	1.00
CHR	NC_010159	Y. pestis	γ proteobacteria	1.00	RECE	NC-010623	B. phymatum	β proteobacteria	0.47
CHR	NC 017154	1. pestis	γproteobacteria	1.00	PL BECE	NC-010625	B. pnymatum	βproteobacteria	0.43
PL	NC_017156	Y. pestis	γproteobacteria	1.00	PL	NC_010679	B. phytofirmans B. phytofirmans	β proteobacteria	0.00
\mathbf{PL}	NC_017159	Y. pestis	γ proteobacteria	1.00	CHR	NC_010681	B. phytofirmans	β proteobacteria	1.00
CHR	NC_017160	Y. pestis	γ proteobacteria	1.00	CHR	NC_006350	B. pseudomallei	β proteobacteria	1.00
CHR	NC_017168	Y. pestis	γ proteobacteria	1.00	RECE	NC_006351	B. pseudomallei	β proteobacteria	1.00
PL	NC_017170	Y. pestis	γ proteobacteria	1.00	CHR	NC_007434	B. pseudomallei	β proteobacteria	1.00
CHR	NC 006155	1. pestis V pseudotubercalogia	γproteobacteria	1.00	CHP	NC 009074	B. pseudomallei	βproteobacteria	1.00
CHR	NC 009708	Y nseudotuberculosis	oproteobacteria	1.00	BECE	NC 009074	B. pseudomallei	Bproteobacteria	0.00
CHR	NC_010465	Y. pseudotuberculosis	γproteobacteria	1.00	CHR	NC_009076	B. pseudomallei	β proteobacteria	1.00
CHR	NC_010634	Y. pseudotuberculosis	γ proteobacteria	1.00	RECE	NC_009078	B. pseudomallei	β proteobacteria	0.00
	-				CHR	NC_012695	B. pseudomallei	β proteobacteria	1.00
cluster :	3 stability me	asure : 0.594261			CHR	NC_017831	B. pseudomallei	β proteobacteria	1.00
PL	NC_007927	Streptomyces sp.	Actinobacteridae	0.43	RECE	NC-017832	B. pseudomallei	β proteobacteria	0.47
PL DI	NC_014919	A. brasilense	αproteobacteria	0.00	BECE	NC_018527	D. pseudomallei	βproteobacteria	0.64
PL	NC.014818	A. lipoferum	aproteobacteria	0.00	CHR	NC_014722	B. rhizoxinica	β proteobacteria	1.00
RECE	NC_011983	A. radiobacter	αproteobacteria	0.00	RECE*	NC_007509	Burkholderia sp.	β proteobacteria	0.27
\mathbf{PL}	NC_011887	M. nodulans	α proteobacteria	0.03	\mathbf{CHR}	NC_007510	<i>B.</i> sp.	β proteobacteria	1.00
\mathbf{PL}	NC_010510	$M.\ radiotolerans$	α proteobacteria	0.36	RECE	NC_007511	Burkholderia sp.	β proteobacteria	0.43
PL	NC_011368	R. leguminosarum	α proteobacteria	0.00	CHR	NC_014117	B. sp.	β proteobacteria	1.00
PĹ DI	NC_006569	R. pomeroyi	α proteobacteria	0.17	RECE	NC_014118	Burkholderia sp.	β proteobacteria	0.27
PL PL	NC 016815	5. jrean S fredii	αproteobacteria	0.00	CHB	NC 014539	Burkholderia sp.	βproteobacteria	1.00
PL	NC_017957	T. mobilis	aproteobacteria	0.00	RECE	NC_014540	Burkholderia sp.	β proteobacteria	0.47
				1.					

01110	NC_015136	Burkholderia sp.	β proteobacteria	1.00	CHR	NC_013851	A. vinosum	γ proteobacteria	0.29
RECE	NC_015137	Burkholderia sp.	β proteobacteria	0.47	CHR	NC_002971	C. burnetii	γ proteobacteria	0.00
CHR	NC_016589	Burkholderia sp.	β proteobacteria	1.00	CHR	NC_009727	C. burnetii	γ proteobacteria	0.36
RECE*	NC_016590	Burkholderia sp.	β proteobacteria	1.00	CHR	NC_010117	C. burnetii	γ proteobacteria	0.50
RECE	NC_016625	Burkholderia sp.	β proteobacteria	0.27	CHR	NC_011527	C. burnetii	γ proteobacteria	1.00
PL	NC 016626	Burkholderia sp.	βproteobacteria	0.00	CHR	NC 011528	C. hurnetii	vproteobacteria	0.50
CHB	NC 017920	Burkholderia sp	Broteobacteria	1.00	CHB	NC 010995	C ignonicus	oproteobacteria	0.14
BECE	NC 017921	Burkholderia sp.	Broteobacteria	0.64	CHR	NC 007963	C. saleriaens	oproteobacteria	0.00
DECE*	NC 017022	Durkholdenia sp.	Proteobacteria	0.04	CHD	NC 018607	C. suierigens	proteobacteria	0.00
RECE*	NC_017922	Burknolaeria sp.	pproteobacteria	0.43	CHR	NC_018697	Cycloclasticus sp.	γ proteobacteria	0.18
RECE	NC_007650	B. thailandensis	proteobacteria	0.43	PL	NC_018985	E. amylovora	γ proteobacteria	0.00
CHR	NC_007651	B. thailandensis	β proteobacteria	1.00	PL	NC_019081	E. coli	γ proteobacteria	0.00
RECE*	NC_009254	$B. \ vietnamiens is$	β proteobacteria	0.88	CHR	NC_017033	F. aurantia	γ proteobacteria	0.33
RECE	NC_009255	B. vietnamiensis	β proteobacteria	1.00	CHR	NC_014366	γ proteobacteriaacterium	γ proteobacteria	0.10
CHR	NC_009256	B. vietnamiensis	β proteobacteria	1.00	CHR	NC_007645	H. chejuensis	γ proteobacteria	0.23
CHR	NC_007951	B. xenovorans	β proteobacteria	1.00	CHR	NC_014532	H. elongata	γ proteobacteria	0.50
RECE	NC 007952	B. xenovorans	βproteobacteria	0.64	CHR	NC 008789	H. halophila	vproteobacteria	0.45
BECE*	NC 007953	B renovorans	Brroteobacteria	0.36	CHB	NC 013422	H neanolitanus	oproteobacteria	0.18
CHB	NC 012104	Cand Accumulibrator	Brioteobacteria	1.00	CHB	NC 017506	M. adhaorana	proteobacteria	0.10
CHR	NC 01594	Cunu. Accumundacter	<i>p</i> proteobacteria	1.00	CHD	NC 016112	M. dunderens	yproteobacteria	0.33
CHR	NC_015856	C. fungivorans	pproteobacteria	1.00	CHR	NC_016112	M. alcaliphilum	γproteobacteria	0.00
CHR	NC_007973	C. metallidurans	β proteobacteria	1.00	CHR	NC_008740	M. aquaeolei	γ proteobacteria	0.20
PL	NC_007974	C. metallidurans	β proteobacteria	0.64	CHR	NC_002977	M. capsulatus	γ proteobacteria	0.64
RECE	NC_015723	$C. \ necator$	β proteobacteria	0.64	CHR	NC_017067	M. hydrocarbonoclasticus	γ proteobacteria	0.50
CHR	NC_015726	C. necator	β proteobacteria	1.00	CHR	NC_015276	M. mediterranea	γ proteobacteria	0.12
CHR	NC_010528	$C. \ taiwanensis$	β proteobacteria	1.00	CHR	NC_015572	M. methanica	γ proteobacteria	0.33
RECE	NC_010530	$C. \ taiwanensis$	β proteobacteria	0.64	CHR	NC_015559	M. posidonica	γ proteobacteria	0.07
CHR	NC_013446	C. testosteroni	βproteobacteria	nan	CHR	NC_018268	Marinobacter sp.	vproteobacteria	0.50
CHB	NC 005085	C violaceum	βproteobacteria	0.95	CHB	NC 009654	Marinomonas sp	oproteobacteria	0.05
CUP	NC 010002	D. acidoworana	Brioteobacteria	1.00	CHB	NC 017857	Mathylophaga sp.	aprotochastoria	0.00
CHR	NC_010002	D. aciaovorans	pproteobacteria	1.00	CIIR	NC_017857	Meingiophaga sp.	γproteobacteria	0.00
CHR	NC_007298	D. aromatica	pproteobacteria	1.00	CHR	NC_013960	N. naiopniius	γproteobacteria	0.50
CHR	NC_015563	Deftia sp.	β proteobacteria	1.00	CHR	NC_007484	N. oceani	γ proteobacteria	0.40
CHR	NC_016616	D. suillum	β proteobacteria	1.00	CHR	NC_014315	N. watsonii	γ proteobacteria	0.18
CHR	NC_014394	G. capsiferriformans	β proteobacteria	1.00	CHR	NC_002516	P. aeruginosa	γ proteobacteria	0.47
CHR	NC_009138	H. arsenicoxydans	β proteobacteria	1.00	CHR	NC_008463	P. aeruginosa	γ proteobacteria	0.64
CHR	NC_014323	H. seropedicae	βproteobacteria	1.00	CHR	NC_009656	P. aeruginosa	vproteobacteria	0.33
CHB	NC 009659	Janthinobacterium sp	βproteobacteria	1.00	CHB	NC 011770	P aeruainosa	vproteobacteria	0.64
CHR	NC 010524	L cholodnii	Broteobacteria	1.00	CHR	NC 017548	P aeruginosa	oproteobacteria	0.50
CUD	NG 010550		pproteobacteria	1.00	CUD	NC 017540	D	proteobacteria	0.30
CHR	NC_012559	L. nongkongensis	pproteobacteria	1.00	CHR	NC_017549	P. aeruginosa	γ proteobacteria	0.43
CHR	NC_007947	M. flagellatus	β proteobacteria	1.00	CHR	NC_018080	P. aeruginosa	γ proteobacteria	0.64
CHR	NC_012969	M. glucosetrophus	β proteobacteria	1.00	CHR	NC_015379	P. brassicacearum	γ proteobacteria	0.00
CHR	NC_012968	M. mobilis	β proteobacteria	1.00	CHR	NC_008027	P. entomophila	γ proteobacteria	0.47
CHR	NC_008825	M. petroleiphilum	β proteobacteria	1.00	CHR	NC_007492	P. fluorescens	γ proteobacteria	0.00
\mathbf{PL}	NC_008826	M. petroleiphilum	β proteobacteria	0.73	CHR	NC_012660	P. fluorescens	γ proteobacteria	0.27
CHR	NC_014733	Methylovorus sp.	βproteobacteria	1.00	CHR	NC_016830	P. fluorescens	vproteobacteria	0.64
CHR	NC 014207	M. versatilis	βproteobacteria	1.00	CHR	NC 017911	P. fluorescens	vproteobacteria	0.64
CHR	NC 004757	N europaea	Broteobacteria	1.00	CHB	NC 015556	P fulua	oproteobacteria	0.27
CHB	NC 008244	N. europaea	Pproteobacteria	1.00	CHD	NC 000420	D m m d min n	proteobacteria	0.21
CHR	NC_008344	N. eutropha	pproteopacteria	1.00	CHR	INC_009439	P. menaocina	γ proteobacteria	0.64
('HR		37 1.10 1	0 1 1 1	1 00	CUID	NG 015410	D I :		0 50
onn	NC_007014	N. multiformis	β proteobacteria	1.00	\mathbf{CHR}	NC_015410	P. mendocina	γ proteobacteria	0.50
CHR	NC_007014 NC_015222	N. multiformis Nitrosomonas sp.	etaproteobacteria etaproteobacteria	1.00 1.00	CHR RECE	NC_015410 NC_006371	P. mendocina P. profundum	γ proteobacteria γ proteobacteria	0.50 0.04
CHR CHR	NC_015222 NC_015731	N. multiformis Nitrosomonas sp. Nitrosomonas sp.	etaproteobacteria etaproteobacteria etaproteobacteria	1.00 1.00 1.00	CHR RECE CHR	NC_015410 NC_006371 NC_004129	P. mendocina P. profundum P. protegens	γ proteobacteria γ proteobacteria γ proteobacteria	0.50 0.04 0.47
CHR CHR CHR	NC_007014 NC_015222 NC_015731 NC_008781	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans	β proteobacteria β proteobacteria β proteobacteria β proteobacteria	1.00 1.00 1.00 nan	CHR RECE CHR CHR	NC_015410 NC_006371 NC_004129 NC_002947	P. mendocina P. profundum P. protegens P. putida	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	0.50 0.04 0.47 0.50
CHR CHR CHR CHR	NC_007014 NC_015222 NC_015731 NC_008781 NC_009379	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius	β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria	1.00 1.00 1.00 nan 1.00	CHR RECE CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_002947 NC_009512	P. mendocina P. profundum P. protegens P. putida P. putida	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.50 0.04 0.47 0.50 0.64
CHR CHR CHR CHR CHR	NC_007014 NC_015222 NC_015731 NC_008781 NC_009379 NC_010531	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius P. necessarius	β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria	1.00 1.00 1.00 nan 1.00 1.00	CHR RECE CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_002947 NC_009512 NC_010322	P. mendocina P. profundum P. protegens P. putida P. putida P. putida	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.50 \\ \hline 0.04 \\ 0.47 \\ 0.50 \\ 0.64 \\ 1.00 \end{array}$
CHR CHR CHR CHR CHR CHR	NC-007944 NC-015222 NC-015731 NC-008781 NC-009379 NC-010531 NC 007948	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius P. necessarius Polaromonas sp.	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 nan 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_002947 NC_009512 NC_010322 NC_010501	P. mendocina P. profundum P. protegens P. putida P. putida P. putida P. putida	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.50 \\ \hline 0.04 \\ 0.47 \\ 0.50 \\ 0.64 \\ 1.00 \\ 0.64 \end{array}$
CHR CHR CHR CHR CHR CHR CHR	NC_007014 NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius P. necessarius Polaromonas sp. Pusillomonas sp.	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 nan 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_002947 NC_009512 NC_010322 NC_010501 NC_015733	P. mendocina P. protegens P. putida P. putida P. putida P. putida P. putida P. putida P. putida	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.50 \\ \hline 0.04 \\ 0.47 \\ 0.50 \\ 0.64 \\ 1.00 \\ 0.64 \\ 0.27 \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR	NC_0075222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp.	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 nan 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_002947 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530	P. mendocina P. profundum P. protegens P. putida P. putida P. putida P. putida P. putida P. putida P. putida	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.50 \\ \hline 0.04 \\ 0.47 \\ 0.50 \\ 0.64 \\ 1.00 \\ 0.64 \\ 0.27 \\ 1.00 \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR PL CHR	NC_0075222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015459	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp.	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530	P. mendocina P. profundum P. protegens P. putida P. putida P. putida P. putida P. putida P. putida P. putida P. putida	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 0.27 1.00 0.64
CHR CHR CHR CHR CHR CHR CHR PL CHR CHR	NC_015242 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015455 NC_015455 NC_015459	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp.	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009129 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017986	P. mendocina P. protegens P. putida P. putida P. putida P. putida P. putida P. putida P. putida P. putida P. putida P. putida	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 0.27 1.00 0.64
CHR CHR CHR CHR CHR CHR CHR PL CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusullomonas sp. Pseudogulbenkiania sp. R. eutropha	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_009512 NC_010322 NC_010501 NC_01533 NC_017530 NC_017530 NC_017986 NC_018220	P. mendocina P. protegens P. putida P. putida	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 0.27 1.00 0.64 0.27
CHR CHR CHR CHR CHR CHR CHR PL CHR CHR CHR RECE	NC_0015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_007347 NC_007348	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha	$\begin{array}{l} \beta \mathrm{proteobacteria} \\ \beta \mathrm{proteobacteria} \end{array}$	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.47	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017986 NC_018220 NC_003892	P. mendocina P. profundum P. protegens P. putida P. putida	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR PL CHR CHR CHR RECE CHR	NC_005222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015459 NC_015459 NC_016002 NC_007347 NC_007348 NC_008313	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR PL CHR	NC_015410 NC_006371 NC_009129 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017986 NC_018220 NC_003892 NC_016147	P. mendocina P. protegens P. putida P. sp. P. sp. P. sp.	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55 \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR CHR RECE CHR RECE	NC_005222 NC_015222 NC_015731 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_016002 NC_007347 NC_007348 NC_008313 NC_008314	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_017530 NC_017530 NC_017530 NC_017986 NC_018220 NC_003892 NC_016147 NC_009434	P. mendocina P. protegens P. putida P. sp.	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR CHR CHR RECE CHR RECE CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015459 NC_015459 NC_015459 NC_015459 NC_007348 NC_007348 NC_008313 NC_008314 NC_007908	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.47 1.00 0.43 nan	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR PL CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_017330 NC_017986 NC_017986 NC_018220 NC_003892 NC_003892 NC_016147 NC_009434 NC_015740	P. mendocina P. protegens P. putida P. sp. M. sp. P. spadix P. stutzeri P. stutzeri	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR CHR RECE CHR RECE CHR CHR	NC_001522 NC_015731 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_007347 NC_007348 NC_007348 NC_008313 NC_008314 NC_007908 NC_017075	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR PL CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017986 NC_018220 NC_003892 NC_008434 NC_017532	P. mendocina P. protegens P. putida P. sputida P. sputida P. sp. P. spodix P. stutzeri P. stutzeri	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.00\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR RECE CHR RECE CHR RECE	NC_001522 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007348 NC_008313 NC_008314 NC_007908 NC_017075 NC_010678	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 nan 1.00 0.64	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_017530 NC_017530 NC_017530 NC_017986 NC_018220 NC_003892 NC_016147 NC_009434 NC_015740 NC_018028	P. mendocina P. protegens P. protegens P. putida P. sputida P. spodix P. stutzeri P. stutzeri P. stutzeri P. stutzeri	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.64\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_016002 NC_007347 NC_007348 NC_007348 NC_008314 NC_007908 NC_017075 NC_010678	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_017530 NC_017530 NC_017986 NC_018220 NC_003892 NC_003892 NC_00344 NC_009434 NC_017532 NC_017532 NC_017532 NC_018177	P. mendocina P. protegens P. putida P. sputida P. sputida P. spolita P. spolita P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri	yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.50\\ 0.00\\ 0.64\\ 0.27\\ 1.00\\ 0.50\\ 0.00\\ 0.00\\ 0.64\\ 0.00\\ 0.0$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR CHR RECE CHR CHR CHR	NC_0015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_016022 NC_007347 NC_007348 NC_008314 NC_007908 NC_017075 NC_010678 NC_010678	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_002947 NC_01322 NC_010501 NC_015733 NC_017530 NC_015733 NC_017530 NC_018220 NC_003892 NC_016147 NC_009434 NC_017532 NC_017532 NC_018028 NC_018177 NC_014924	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputida P. sputida P. sputida P. sputida P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE	NC_0015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_008313 NC_008314 NC_008314 NC_007908 NC_010682 NC_010682 NC_010682 NC_012856	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii	βproteobacteria βproteobacteria	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017580 NC_017580 NC_018220 NC_00842 NC_016147 NC_009434 NC_015740 NC_018028 NC_018177 NC_014924	P. mendocina P. protegens P. protegens P. putida P. sp. P. spadix P. stutzeri P. stutzeri	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.47\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015458 NC_007347 NC_007347 NC_007348 NC_007348 NC_007348 NC_007348 NC_007908 NC_017075 NC_010682 NC_012856 NC_012856	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_017530 NC_017530 NC_017530 NC_017986 NC_018220 NC_003892 NC_016147 NC_009434 NC_017532 NC_017532 NC_018177 NC_014924 NC_004578	P. mendocina P. protegens P. protegens P. putida P. putida P. putida P. putida P. putida P. putida P. putida P. putida P. putida P. sputida P. spolita P. spolita P. stutzeri P. stutzeri	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.57\\ 1.00\\ 0.55\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.47\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR	NC.015222 NC.015221 NC.003731 NC.009379 NC.010531 NC.007948 NC.015458 NC.015459 NC.016002 NC.007347 NC.007348 NC.008313 NC.008314 NC.007908 NC.017075 NC.010678 NC.012856 NC.012856 NC.012857 NC.003295	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum	βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_009512 NC_010322 NC_015733 NC_017530 NC_015733 NC_017530 NC_018220 NC_003892 NC_00892 NC_008434 NC_0017532 NC_018028 NC_018177 NC_014924 NC_004578 NC_005773	P. mendocina P. protegens P. putida P. sputida P. sputida P. sputzeri P. stutzeri P.	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.47\\ 0.00\\ 0.0$
CHR CHR CHR CHR CHR CHR CHR PL CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_008313 NC_008314 NC_008314 NC_007908 NC_010682 NC_010682 NC_012856 NC_012857 NC_003295 NC_003296	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.47 1.00 0.43 nan 1.00 0.64 1.00 0.64	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_015733 NC_017530 NC_017530 NC_017986 NC_017530 NC_017986 NC_017986 NC_01820 NC_008434 NC_016147 NC_009434 NC_015740 NC_017532 NC_018028 NC_018177 NC_014924 NC_005773 NC_007005	P. mendocina P. protegens P. protegens P. putida P. sp. P. spadix P. stutzeri P.	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.40\\ 0.40\\ 0.40\\ 0.00\\ 0.00\\ 0.00\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_007348 NC_007348 NC_008314 NC_007908 NC_017075 NC_010682 NC_010682 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_01533 NC_017530 NC_017530 NC_017530 NC_017986 NC_018200 NC_008392 NC_016147 NC_009434 NC_0017532 NC_0161477 NC_014924 NC_004578 NC_0045773 NC_007005 NC_014258	P. mendocina P. protegens P. protegens P. putida P. spodix P. stutzeri P. stutzeri	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_007347 NC_007348 NC_007348 NC_007347 NC_007348 NC_007343 NC_007348 NC_007908 NC_017075 NC_010678 NC_012856 NC_012856 NC_012857 NC_003295 NC_003295 NC_003295 NC_014310	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_002947 NC_01322 NC_010501 NC_015733 NC_017530 NC_015733 NC_015730 NC_00892 NC_00892 NC_0045740 NC_014528 NC_004578 NC_004578 NC_004578 NC_007005 NC_014258 NC_018868	P. mendocina P. protegens P. putida P. sputida P. sputida P. sputida P. sputida P. sputida P. sputida P. sputida P. stutzeri P. stutzeri S. sagarivorans	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.47\\ 0.00\\ 0.47\\ 0.00\\ 0.00\\ 0.38\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_007348 NC_008313 NC_008314 NC_008313 NC_008314 NC_007908 NC_010682 NC_010682 NC_010682 NC_012856 NC_012856 NC_012857 NC_003295 NC_003296 NC_014307 NC_014311	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017530 NC_017520 NC_00892 NC_016147 NC_009434 NC_015740 NC_017532 NC_018028 NC_018028 NC_018177 NC_014924 NC_004578 NC_005773 NC_007005 NC_014258 NC_018688 NC_007912	P. mendocina P. protegens P. protegens P. putida P. sp. P. spadix P. stutzeri P.	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.40\\ 0.40\\ 0.40\\ 0.40\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.55\\ 0.50\\ 0.5$
CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_007348 NC_007348 NC_008313 NC_008314 NC_007908 NC_017075 NC_010682 NC_010682 NC_012856 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307 NC_014311 NC_017574	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_01533 NC_017530 NC_017530 NC_017530 NC_017986 NC_018220 NC_008392 NC_016147 NC_009434 NC_016147 NC_009434 NC_018740 NC_018177 NC_018028 NC_018177 NC_018177 NC_014924 NC_004578 NC_007005 NC_014258 NC_018868 NC_018868 NC_018500	P. mendocina P. protegens P. protegens P. putida P. sp. P. spadix P. stutzeri P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. sugans S. agarivorans S. degradans	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.55\\ 0.50\\ 0.00\\ 0.55\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.40\\ 0.47\\ 0.00\\ 0.38\\ 0.50\\ 0.0$
CHR CHR CHR CHR CHR CHR CHR PL CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR	NC.015222 NC.015221 NC.015731 NC.009379 NC.010531 NC.007948 NC.015459 NC.015459 NC.015459 NC.016002 NC.007347 NC.007348 NC.008314 NC.007308 NC.017075 NC.012856 NC.012856 NC.012856 NC.012856 NC.012857 NC.003295 NC.003296 NC.014307 NC.014310 NC.014311 NC.017575	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_015733 NC_017530 NC_00892 NC_00892 NC_00892 NC_016147 NC_004574 NC_017532 NC_014028 NC_017532 NC_014028 NC_004578 NC_004578 NC_004578 NC_004258 NC_018868 NC_007912 NC_010500 NC_010943	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputia P. stutaeri P. stutzeri P. stutzeri S. suwonensis P. syringae P. syringae P. vagans S. agarivorans S. degradans S. enterica S. matrophilia	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.47\\ 0.00\\ 0.47\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.40\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR PL CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_007348 NC_008313 NC_008314 NC_008313 NC_008314 NC_008313 NC_008314 NC_007908 NC_010682 NC_010682 NC_010682 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307 NC_014311 NC_017574 NC_017575	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Reutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 nan 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_015733 NC_017530 NC_017530 NC_017530 NC_017530 NC_017532 NC_006147 NC_009434 NC_016147 NC_009434 NC_0161477 NC_0148028 NC_018028 NC_018028 NC_0180273 NC_004578 NC_005773 NC_007005 NC_014258 NC_014258 NC_014258 NC_016147 NC_019943 NC_010710	P. mendocina P. protegens P. protegens P. putida P. spolita P. spolita P. spolita P. spolita P. stutzeri P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae S. agarivorans S. degradans S. maltophilia	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.40\\ 0.40\\ 0.40\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.00\\ 0.67\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_007348 NC_008313 NC_008314 NC_007908 NC_017075 NC_010682 NC_010682 NC_012856 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307 NC_014307 NC_014311 NC_017574 NC_017575 NC_015677 NC_01359	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_01501 NC_017530 NC_017530 NC_017530 NC_017530 NC_017530 NC_017986 NC_018220 NC_008392 NC_016147 NC_009434 NC_0161477 NC_014924 NC_018028 NC_018177 NC_014924 NC_004578 NC_007005 NC_014258 NC_007005 NC_014258 NC_018868 NC_007912 NC_0105407	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputida P. spolita P. spodix P. stutzeri P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. sugans S. adgradans S. enterica S. maltophilia	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.55\\ 0.50\\ 0.00\\ 0.55\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.40\\ 0.67\\ 5.5\\ 0.55$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR CHR PL CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007348 NC.008314 NC.007348 NC.007908 NC.017075 NC.012856 NC.012856 NC.012856 NC.012857 NC.003296 NC.014307 NC.014310 NC.014310 NC.017575 NC.015677 NC.015677 NC.013959	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Reutropha R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_009512 NC_010322 NC_015733 NC_017530 NC_015733 NC_017530 NC_018200 NC_00892 NC_00892 NC_00892 NC_016147 NC_009434 NC_017532 NC_018028 NC_017532 NC_018028 NC_017532 NC_014258 NC_004578 NC_004578 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_019423 NC_0109413 NC_010711 NC_015947	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputiea P. sputzeri P. stutzeri P. stutzeri S. agarivorans S. degradans S. maltophilia S. maltophilia	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.40\\ 0.67\\ 0.55\\ 0.6$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_008313 NC_008314 NC_008313 NC_008314 NC_008314 NC_007908 NC_010682 NC_010682 NC_010682 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307 NC_014311 NC_017575 NC_013959 NC_01643	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017530 NC_017530 NC_018220 NC_008892 NC_016147 NC_009434 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017537 NC_007005 NC_014258 NC_007912 NC_010500 NC_010943 NC_011071 NC_015947 NC_017671 NC_017671	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputa P. sputa P. spadix P. stutzeri P. stutzeri S. sugariorans S. agarivorans S. degradans S. maltophilia S. maltophilia S. maltophilia	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.47\\ 0.00\\ 0.00\\ 0.47\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.67\\ 0.55\\ 0.40\\ 0.67\\ 0.55\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.50\\ 0.55\\ 0.50\\ 0.55\\ 0.50\\ 0.5$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR CHR PL CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_016002 NC_007347 NC_007347 NC_007348 NC_007348 NC_007348 NC_007348 NC_007348 NC_007348 NC_007348 NC_010682 NC_012856 NC_012856 NC_012856 NC_012856 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307 NC_014301 NC_017575 NC_015677 NC_015675 NC_016043 NC_007404	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010501 NC_01733 NC_017530 NC_017530 NC_017530 NC_017530 NC_017530 NC_017986 NC_017532 NC_016147 NC_009434 NC_016147 NC_016328 NC_018028 NC_018028 NC_018073 NC_007005 NC_014258 NC_007005 NC_014258 NC_007912 NC_010500 NC_01943 NC_010941 NC_016947 NC_016947 NC_016711 NC_018889	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputida P. spolita P. spadix P. stutzeri P. stutzeri S. stutzeri P. syringae P. syringae P. vagans S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia T. sp.	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.40\\ 0.40\\ 0.40\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.55\\ 0.40\\ 0.21\\ 0.55\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.25\\ 0.40\\ 0.21\\ 0.25\\ 0.25\\ 0.20\\ 0.2$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.007348 NC.008314 NC.007908 NC.017075 NC.012856 NC.012856 NC.012856 NC.012857 NC.003296 NC.014307 NC.014310 NC.014310 NC.017575 NC.015677 NC.015677 NC.013959 NC.015643 NC.007404 NC.014914	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Reutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. asinigenitalis T. denitrificans T. equigenitalis	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_002947 NC_01322 NC_010501 NC_015733 NC_017530 NC_015733 NC_018200 NC_003892 NC_016147 NC_009434 NC_007432 NC_017532 NC_018028 NC_017532 NC_018028 NC_017532 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_010901 NC_010943 NC_011071 NC_013849 NC_011901	P. mendocina P. protegens P. protegens P. putida P. sputia P. stutaeri P. stutzeri P. stutzeri S. usurenensis P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.57\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.40\\ 0.47\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.40\\ 0.67\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_016002 NC_007347 NC_007347 NC_008313 NC_008313 NC_008314 NC_008314 NC_007908 NC_010682 NC_010682 NC_010682 NC_012856 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307 NC_014311 NC_017575 NC_014311 NC_017575 NC_016677 NC_013959 NC_016043 NC_014914 NC_014108	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. asinigenitalis T. equigenitalis	βproteobacteria βproteobacteri	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017530 NC_017530 NC_018028 NC_00892 NC_016147 NC_009434 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017537 NC_004578 NC_004578 NC_004578 NC_004578 NC_004578 NC_004578 NC_004578 NC_018688 NC_01921 NC_010500 NC_0105047 NC_017671 NC_015947 NC_017671 NC_013889 NC_011001 NC_012997	P. mendocina P. protegens P. protegens P. putida P. sputia P. sputia P. sputia P. stutzeri P. stutzeri S. suvonensis P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae S. agardans S. agardans S. enterica S. maltophilia S. maltophilia S. maltophilia T. sp. T. sulfidophilus T. turnerae	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.47\\ 0.00\\ 0.0$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015222 NC_015731 NC_008781 NC_009379 NC_010531 NC_007948 NC_015458 NC_015458 NC_015459 NC_016002 NC_007347 NC_007347 NC_007348 NC_007348 NC_007348 NC_007348 NC_007348 NC_007348 NC_007348 NC_010682 NC_012856 NC_012856 NC_012856 NC_012857 NC_003296 NC_014307 NC_014301 NC_014311 NC_017574 NC_013599 NC_016043 NC_014043 NC_007404 NC_014108	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017530 NC_017530 NC_017532 NC_016147 NC_009434 NC_016147 NC_009434 NC_016147 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_005773 NC_007005 NC_014258 NC_005773 NC_007005 NC_014258 NC_007912 NC_01943 NC_01947 NC_015947 NC_012997 NC_013722	P. mendocina P. protegens P. protegens P. putida P. sputida P. spodix P. spadix P. stutzeri P. stutzeri S. stutzeri S. matophilia S. maltophilia S. maltophilia S. maltophilia T. sp. T. sulfidophilus T. turnerae X. albilineans	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.40\\ 0.40\\ 0.40\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.44\\ 0.86\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.012856 NC.012856 NC.012856 NC.012857 NC.003295 NC.003295 NC.0014307 NC.014310 NC.014311 NC.017575 NC.015677 NC.015677 NC.015677 NC.015643 NC.007404 NC.014914 NC.014918 NC.01462	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. asinigenitalis T. equigenitalis T. equigenitalis T. intermedia T. sp.	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_002947 NC_01322 NC_010501 NC_015733 NC_017530 NC_017530 NC_018200 NC_003892 NC_016147 NC_009434 NC_017532 NC_018028 NC_017532 NC_018028 NC_017532 NC_014258 NC_011071 NC_015947 NC_013889 NC_011901 NC_012997 NC_013722 NC_003919	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputida P. sputies P. stutzeri P. stutzeri S. usurenensis P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae S. agarivorans S. degradans S. degradans S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. turnerae X. albilineans X. axonopodis	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.40\\ 0.47\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.40\\ 0.67\\ 0.55\\ 0.40\\ 0.67\\ 0.55\\ 0.40\\ 0.41\\ 0.81\\ 0.40\\ 0.44\\ 0.86\\ 1.00\\ 0.00\\ 0.10\\ 0.00\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.44\\ 0.86\\ 0.21\\ 0.40\\ 0.41\\ 0.86\\ 0.21\\ 0.40\\ 0.41\\ 0.86\\ 0.21\\ 0.40\\ 0.41\\ 0.86\\ 0.21\\ 0.40\\ 0.40\\ 0.21\\ 0.40\\ 0.21\\ 0.20\\ 0.2$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.008313 NC.008313 NC.008314 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.012857 NC.003296 NC.014301 NC.014311 NC.017575 NC.01677 NC.013959 NC.016043 NC.014914 NC.014153 NC.014158 NC.014158 NC.014158	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. asinigenitalis T. equigenitalis T. equigenitalis T. intermedia T. sp. T. sp.	βproteobacteria βproteobacteri	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_002947 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017530 NC_018220 NC_003892 NC_016147 NC_003892 NC_016147 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017537 NC_004578 NC_004578 NC_004578 NC_004578 NC_004578 NC_004578 NC_010428 NC_018688 NC_010912 NC_010500 NC_010943 NC_011071 NC_015947 NC_017671 NC_013889 NC_011901 NC_012997 NC_013722 NC_003919 NC_016010	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputia P. stutaeri P. stutzeri P. stutzeri S. suzonensis S. agarivorans S. agarivorans S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. turnerae X. azonopodis X. azonopodis	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.47\\ 0.00\\ 0.0$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.012856 NC.012856 NC.012856 NC.012856 NC.012856 NC.014307 NC.014301 NC.014311 NC.017574 NC.015677 NC.015677 NC.015677 NC.015677 NC.015677 NC.015677 NC.015677 NC.0156139 NC.016043 NC.007404 NC.014914 NC.014153 NC.014153 NC.014145	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017530 NC_017530 NC_017532 NC_016147 NC_009434 NC_016147 NC_009434 NC_016147 NC_01618028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_01943 NC_01943 NC_01943 NC_01947 NC_015947 NC_013889 NC_013889 NC_011901 NC_012997 NC_013722 NC_003902	P. mendocina P. protegens P. protegens P. putida P. sputida P. spo S. spo P. spadix P. stutzeri P. stutzeri S. stutzeri S. matophilia S. maltophilia S. maltophilia	yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 1.00 0.57 1.00 0.55 0.50 0.00 0.64 1.00 0.55 0.50 0.00 0.64 1.00 0.40 0.55 0.50 0.00 0.64 1.00 0.55 0.50 0.00 0.64 1.00 0.55 0.50 0.00 0.64 1.00 0.55 0.50 0.00 0.40
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.012856 NC.012857 NC.016825 NC.013295 NC.014307 NC.014310 NC.014311 NC.017574 NC.01577 NC.015677 NC.015677 NC.015677 NC.015643 NC.007404 NC.014153 NC.01462 NC.014153 NC.01462 NC.018761 NC.018108	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Reutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. asinigenitalis T. denitrificans T. equigenitalis T. equigenitalis T. intermedia T. sp. T. sp. V. eiseniae V naradocrus	βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_002947 NC_01322 NC_010501 NC_015733 NC_017530 NC_017530 NC_018200 NC_003892 NC_016147 NC_009434 NC_007434 NC_017532 NC_018028 NC_017532 NC_018028 NC_017532 NC_014258 NC_011071 NC_015947 NC_013889 NC_011901 NC_012997 NC_013722 NC_003912 NC_003902 NC_00786	P. mendocina P. protegens P. protegens P. putida P. sputia P. stutaeri P. stutzeri P. stutzeri S. usurenensis P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae S. agarivorans S. degradans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. turnerae X. albilineans X. axonopodis X. campestris	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 0.27\\ 1.00\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.40\\ 0.67\\ 0.55\\ 0.40\\ 0.67\\ 0.55\\ 0.40\\ 0.40\\ 0.67\\ 0.55\\ 0.40\\ 0.40\\ 0.44\\ 0.86\\ 1.00\\ 0.40\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ 0.55\\ 0.40\\ 0.21\\ 0.55\\ 0.40\\ 0.21\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.40\\ 0.55\\ 0.55\\ 0.40\\ 0.55\\ 0.55\\ 0.40\\ 0.55\\ 0.55\\ 0.55\\ 0.55\\ 0.40\\ 0.55\\ 0.5$
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.008313 NC.008313 NC.008314 NC.008313 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.012856 NC.014301 NC.014311 NC.017575 NC.014311 NC.017575 NC.016677 NC.013959 NC.016043 NC.014914 NC.014914 NC.014153 NC.014153 NC.014153 NC.014153 NC.014153 NC.014153 NC.014154 NC.014153 NC.01453	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_004129 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017530 NC_017530 NC_018220 NC_00892 NC_016147 NC_009434 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017537 NC_004578 NC_004578 NC_004578 NC_004578 NC_0104258 NC_010500 NC_010500 NC_0105947 NC_017671 NC_013889 NC_011901 NC_012997 NC_016010 NC_003902 NC_007568	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputia P. stutaeri P. stutzeri P. stutzeri S. suzonensis S. agarivorans S. agarivorans S. degradans S. enterica S. maltophilia S. campestris X. campestris X. campestris	yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.55 0.50 0.50 0.50 0.50 0.50 0.60 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.55 0.50 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.64 1.00 0.64 0.27 0.55 0.50 0.50 0.50 0.50 0.00 0.40
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.008314 NC.007348 NC.010682 NC.010682 NC.012856 NC.012856 NC.012856 NC.012857 NC.003296 NC.014301 NC.014301 NC.014301 NC.014301 NC.014301 NC.014914 NC.014153 NC.016622 NC.014153 NC.016622 NC.014153 NC.012791 NC.012791 NC.012791 NC.012791	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017986 NC_017530 NC_017986 NC_017986 NC_018200 NC_00892 NC_016147 NC_009434 NC_015740 NC_017532 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_018028 NC_007005 NC_014258 NC_005773 NC_007005 NC_014258 NC_018688 NC_007912 NC_010500 NC_01943 NC_011947 NC_015947 NC_015947 NC_013889 NC_013722 NC_003919 NC_013722 NC_003919 NC_013722 NC_003902 NC_007086 NC_007508	P. mendocina P. protegens P. protegens P. putida P. spolita P. spolita P. spolita P. stutzeri P. stutzeri S. stutzeri S. sustonensis P. syringae P. syringae P. syringae P. syringae P. syringae P. syringae S. agarivorans S. degradans S. degradans S. maltophilia S. turnerae X. axonopodis X. campestris X. campestris X. campestris	yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 1.00 0.57 1.00 0.55 0.50 0.00 0.64 1.00 0.55 0.50 0.00 0.64 1.00 0.40
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR PL CHR PL CHR PL CHR PL CHR PL CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.012857 NC.01682 NC.014307 NC.014301 NC.014574 NC.017575 NC.016073 NC.016073 NC.016073 NC.016073 NC.016043 NC.014153 NC.014153 NC.014153 NC.014155 NC.003296 NC.014155 NC.018791 NC.012791 NC.012792 NC.014931	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. asinigenitalis T. denitrificans T. equigenitalis T. equigenitalis T. intermedia T. sp. T. sp. V. eiseniae V. paradoxus V. paradoxus	βproteobacteria βproteobacteri	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_009512 NC_010322 NC_015733 NC_017530 NC_015733 NC_017530 NC_018220 NC_003892 NC_016147 NC_009434 NC_007532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_016010 NC_013947 NC_013889 NC_011901 NC_013947 NC_013947 NC_013947 NC_013947 NC_013947 NC_013919 NC_016010 NC_003902 NC_007508 NC_01688	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputia P. stutzeri P. stutzeri S. agarivorans S. degradans S. degradans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. turnerae X. albilineans X. campestris X. campestris X. campestris X. campestris X. campestris X. campestris	yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.55 0.50 0.50 0.50 0.50 0.50 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.55 0.50 0.50 0.50 0.50 0.55 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.40 0.55 0.50 0.50 0.50 0.00 0.00 0.00 0.40
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008313 NC.008313 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.012857 NC.003295 NC.003295 NC.014301 NC.014311 NC.017575 NC.013959 NC.016043 NC.014914 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014914 NC.014931 NC.014931 NC.014931 NC.008260	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum T. solanacearum R. solanacearum	βproteobacteria βproteobacteri	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_00912 NC_009512 NC_010322 NC_010501 NC_015733 NC_017530 NC_017530 NC_017530 NC_018220 NC_00892 NC_016147 NC_009434 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_010500 NC_010500 NC_010943 NC_011071 NC_015947 NC_017671 NC_013889 NC_011901 NC_012997 NC_016010 NC_007908 NC_007086 NC_007688 NC_007688 NC_007688 NC_007688	P. mendocina P. protegens P. protegens P. putida P. sputida P. sputia P. stutzeri P. stutzeri S. suzonensis S. agarivorans S. agarivorans S. degradans S. enterica S. maltophilia S. campestris X. campestris X. campestris X. campestris X. campestris X. campestris X. campestris X. campestris X. campestris	 yproteobacteria ypr	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.55 0.50 0.50 0.50 0.50 0.60 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.55 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.64 1.00 0.55 0.50 0.50 0.50 0.50 0.55 0.50 0.50 0.64 1.00 0.55 0.50 0.50 0.50 0.50 0.64 1.00 0.55 0.50 0.50 0.50 0.64 1.00 0.00 0.41 0.40 0.44 0.40 0.55 0
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.008314 NC.007348 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.012857 NC.003296 NC.014301 NC.014301 NC.014301 NC.014301 NC.014301 NC.014914 NC.014914 NC.014153 NC.016623 NC.014153 NC.014153 NC.014145 NC.012791 NC.012792 NC.012791 NC.012792 NC.012850	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. solanacearum R. sola	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_002947 NC_01322 NC_015733 NC_017530 NC_015733 NC_017530 NC_015730 NC_003892 NC_003892 NC_016147 NC_004578 NC_015740 NC_015740 NC_015740 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014259 NC_014259 NC_014259 NC_017671 NC_013722 NC_013722 NC_003902 NC_007508 NC_007508 NC_007508 NC_007508	P. mendocina P. protegens P. putida P. sputida P. sputida P. sputieri P. stutzeri P. stutzeri S. sutceri S. susteria S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia S. maltophilia T. sp. T. sulfidophilus T. turnerae X. axonopodis X. campestris X. campestris	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.55\\ 0.50\\ 0.00\\ 0.55\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.40\\ 0.40\\ 0.67\\ 0.55\\ 0.40\\ 0.21\\ 0.40\\ 0.40\\ 0.55\\ 1.00\\ 0.40\\ 0.55\\ 1.00\\ 0.40\\ 0.55\\ 0.18\\ \end{array}$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.014307 NC.014307 NC.014301 NC.017574 NC.017575 NC.016073 NC.016073 NC.016075 NC.016075 NC.016073 NC.016073 NC.016043 NC.014153 NC.014153 NC.014153 NC.014153 NC.014153 NC.014914 NC.014791 NC.012791 NC.012792 NC.014931 NC.008850 NC.015851	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. aequigenitalis T. equigenitalis T. equigenitalis T. intermedia T. sp. T. sp. T. sp. V. eiseniae V. paradoxus V. paradoxus A. borkumensis A. caldus	βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_009512 NC_010322 NC_015733 NC_017530 NC_015733 NC_017530 NC_018200 NC_00892 NC_00892 NC_016147 NC_009434 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_010943 NC_010943 NC_010943 NC_010943 NC_010943 NC_010943 NC_010943 NC_010943 NC_011071 NC_013889 NC_011071 NC_013889 NC_011901 NC_013947 NC_013889 NC_011901 NC_013919 NC_013919 NC_003919 NC_016010 NC_003902 NC_007508 NC_016888 NC_007508 NC_016888 NC_007217 NC_002488 NC_004556	 P. mendocina P. protegens P. protida P. putida P. sutida P. spadix P. stutzeri S. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri P. stutzeri S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia T. sulfidophilus T. turnerae X. albilineans X. campestris X. campestris X. campestris X. campestris X. fastidiosa X. fastidiosa 	yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.55 0.50 0.50 0.50 0.50 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.50 0.00 0.40 0.55 0.40 0.21 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.55 0.40 0.40 0.40 0.55 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.18 0.50
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015459 NC.016002 NC.007347 NC.007347 NC.008313 NC.008313 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.012856 NC.014301 NC.014311 NC.014311 NC.014311 NC.014311 NC.014311 NC.014311 NC.014311 NC.014311 NC.014313 NC.014914 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014153 NC.014914 NC.014914 NC.014914 NC.014914 NC.014914 NC.014914 NC.014914 NC.014914 NC.014931 NC.014931 NC.008260 NC.015851 NC.018691	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. sol	βproteobacteria βproteobacteri	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017530 NC_017530 NC_018220 NC_003892 NC_016147 NC_009434 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_014224 NC_004578 NC_010500 NC_010943 NC_011071 NC_015947 NC_017671 NC_013889 NC_011901 NC_012997 NC_013722 NC_003919 NC_016010 NC_007508 NC_017271 NC_002488 NC_017271 NC_002488 NC_010513	 P. mendocina P. protegens P. putida P. sutida P. spadix P. stutzeri S. suonensis P. syringae P. syringae P. syringae P. syringae P. syringae S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. turnerae X. akonopodis X. campestris X. fastidiosa X. fastidiosa X. fastidiosa X. fastidiosa 	 yproteobacteria ypr	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.55 0.50 0.50 0.50 0.50 0.60 0.64 1.00 0.47 0.00 0.64 1.00 0.47 0.55 0.50 0.00 0.64 1.00 0.47 0.55 0.50 0.50 0.50 0.50 0.50 0.55 0.50 0.64 1.00 0.55 0.50 0.50 0.50 0.55 0.50 0.50 0.64 1.00 0.55 0.50 0.50 0.55 0.50 0.50 0.64 1.00 0.47 0.55 0.50 0.50 0.64 1.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.55 0.50 0.55 0.40 0.40 0.44 0.27 0.55 0.40 0.40 0.44 0.55 0.55 0.50 0.50 0.00 0.47 0.00 0.47 0.00 0.47 0.55 0.55 0.50 0.55 0.50 0.00 0.47 0.40 0.47 0.55 0.40 0.40 0.47 0.55 0.55 0.40 0.40 0.47 0.55 0.40 0.40 0.47 0.55 0.40 0.40 0.47 0.55 0.40 0.40 0.47 0.40 0.55 0.55 0.33
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.012856 NC.012856 NC.012856 NC.014301 NC.014301 NC.014301 NC.014301 NC.014301 NC.01403 NC.01403 NC.01403 NC.014153 NC.016623 NC.014153 NC.016623 NC.014153 NC.016826 NC.012791 NC.012792 NC.012792 NC.012792 NC.012850 NC.012850 NC.018851 NC.008260 NC.018851 NC.018691 NC.008340	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pseudogulbenkiania sp. R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. solanacearum V. pickettii T. equigenitalis T. equigenitalis T. equigenitalis T. intermedia T. sp. T. sp. V. eiseniae V. paradoxus V. paradoxus A. borkumensis A. caldus A. caldus A. chrischii	β proteobacteria β proteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_002947 NC_002947 NC_002947 NC_01322 NC_015733 NC_017530 NC_015733 NC_015740 NC_003892 NC_016147 NC_0045740 NC_015740 NC_015740 NC_015740 NC_015747 NC_014228 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_010943 NC_010943 NC_010943 NC_017671 NC_013722 NC_013722 NC_013722 NC_003902 NC_007508 NC_007508 NC_017518 NC_007518 NC_007518 NC_007513 NC_010517 NC_010577	 P. mendocina P. protegens P. putida P. suticari P. stutzeri S. stateri S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia T. sulfidophilus T. turnerae X. axonopodis X. campestris X. campestris X. campestris X. campestris X. campestris X. fastidiosa X. fastidiosa X. fastidiosa 	yproteobacteria yproteobacteria	$\begin{array}{c} 0.50\\ \hline 0.04\\ 0.47\\ 0.50\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.55\\ 0.50\\ 0.00\\ 0.55\\ 0.50\\ 0.00\\ 0.64\\ 1.00\\ 0.40\\ 0.47\\ 0.00\\ 0.00\\ 0.40\\ 0.47\\ 0.00\\ 0.00\\ 0.38\\ 0.50\\ 0.00\\ 0.40\\ 0.40\\ 0.40\\ 0.41\\ 0.40\\ 0.55\\ 1.00\\ 0.40\\ 0.55\\ 0.18\\ 0.50\\ 0.40\\ 0.40\\ 0.40\\ 0.40\\ 0.40\\ 0.40\\ 0.40\\ 0.55\\ 0.18\\ 0.50\\ 0.4$
CHR CHR CHR CHR CHR CHR CHR CHR CHR RECE CHR RECE CHR RECE CHR RECE CHR PL CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.008781 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.016002 NC.007347 NC.007347 NC.007348 NC.008314 NC.008314 NC.008314 NC.008314 NC.008314 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.014307 NC.014307 NC.014307 NC.014301 NC.017574 NC.017575 NC.016043 NC.016043 NC.014053 NC.014153 NC.014153 NC.014153 NC.014153 NC.014153 NC.014153 NC.014291 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC.01441 NC	N. multiformis Nitrosomonas sp. Nitrosomonas sp. P. naphthalenivorans P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. aequigenitalis T. equigenitalis T. equigenitalis T. equigenitalis T. equigenitalis T. intermedia T. sp. T. sp. V. eiseniae V. paradoxus V. paradoxus A. borkumensis A. caldus A. caldus A. caldus A. caldus A. cehrichii A. ferrivorans	βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.47 1.00 0.47 1.00 0.43 nan 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.64 1.00 0.50 1.00 1.00 1.00 0.64 1.00 0.50 1.00 1.00 1.00 0.64 1.00 0.33 1.00 1.00 1.00 1.00 0.64 1.00 0.03 1.00 1.00 1.00 1.00 0.64 1.00 0.50 1.00 1.00 1.00 0.64 1.00 0.64 1.00 0.33 1.00 1.00 1.00 1.00 1.00 0.64 1.00 0.33 1.00 1.00 1.00 1.00 1.00 0.64 1.00 0.33 1.00 1.00 1.00 1.00 1.00 0.64 1.00 0.03 1.00 1.00 1.00 1.00 1.00 0.64 1.00 1.00 1.00 0.64 1.00 0.03 1.00 1.00 1.00 1.00 1.00 0.03 1.00 1	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_002947 NC_009512 NC_010322 NC_015733 NC_017530 NC_015733 NC_017530 NC_015733 NC_018200 NC_003892 NC_016147 NC_009434 NC_017532 NC_018028 NC_017532 NC_017532 NC_017532 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_014258 NC_010943 NC_010943 NC_010943 NC_011071 NC_013947 NC_013889 NC_011901 NC_013947 NC_013889 NC_011901 NC_013947 NC_013947 NC_013889 NC_011901 NC_013947 NC_013947 NC_013947 NC_013948 NC_007508 NC_007508 NC_016888 NC_007508 NC_016888 NC_007518 NC_01658 NC_010513 NC_017562	 P. mendocina P. protegens P. protida P. putida P. sutida P. spadix P. stutzeri S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia S. maltophilia X. axonopodis X. campestris X. campestris X. campestris X. campestris X. fastidiosa X. fastidiosa X. fastidiosa X. fastidiosa X. fastidiosa X. fastidiosa 	yproteobacteria yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.55 0.50 0.50 0.50 0.50 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.40 0.40 0.55 0.50 0.40 0.55 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.50 0.50 0.50 0.50 0.50 0.00 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.50 0.50 0.50 0.00 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.40 0.55 0.40 0.40 0.40 0.55 0.40 0.40 0.55 0.40 0.21 0.40 0.40 0.40 0.40 0.40 0.55 0.40 0.21 0.40 0.40 0.40 0.40 0.40 0.55 0.40 0.40 0.55 0.40 0.55 0.18 0.50 0.33 0.40 0.22 0.33 0.40 0.22 0.33 0.40 0.22 0.33 0.40 0.22 0.33 0.40 0.22 0.33 0.40 0.22 0.22 0.25 0.33 0.40 0.22 0.25 0.33 0.40 0.22 0.25 0.33 0.40 0.22 0.25 0.33 0.40 0.22 0.55 0.33 0.40 0.22 0.55
CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.015222 NC.015221 NC.015731 NC.009379 NC.010531 NC.007948 NC.015458 NC.015458 NC.015459 NC.016002 NC.007347 NC.007347 NC.007348 NC.008313 NC.008314 NC.008314 NC.007908 NC.010682 NC.010682 NC.010682 NC.012856 NC.012856 NC.014301 NC.014311 NC.014310 NC.014311 NC.014311 NC.017575 NC.016043 NC.014311 NC.014311 NC.014311 NC.014311 NC.014313 NC.014044 NC.014044 NC.014153 NC.014044 NC.014153 NC.014044 NC.014153 NC.014044 NC.014153 NC.014031 NC.014311 NC.014311 NC.014313 NC.014914 NC.014153 NC.014931 NC.014931 NC.008260 NC.015851 NC.018691 NC.008340 NC.01560	N. multiformis Nitrosomonas sp. P. naphthalenivorans P. necessarius P. necessarius Polaromonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. Pusillomonas sp. R. eutropha R. eutropha R. eutropha R. eutropha R. eutropha R. ferrireducens R. gelatinosus R. pickettii R. pickettii R. pickettii R. pickettii R. solanacearum R. tataouinensis S. lithotrophicus T. asinigenitalis T. equigenitalis T. equigenitalis T. equigenitalis T. equigenitalis T. equigenitalis T. sp. Y. paradoxus V. paradoxus V. paradoxus A. borkumensis A. caldus A. caldus A. dieselolei A. ehrlichii A. ferrivorans A. vinelandii	βproteobacteria βproteobacteri	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.64 1.00 0.43 1.00 0.64 1.00 0.43 1.00 0.64 1.00 0.43 1.00 0.64 1.00 0.43 1.00 0.43 1.00 0.64 1.00 0.50 1.00 0.50 1.00 0.50 1.00 0.50 1.00 0.50 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.64 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.43 1.00 0.50 1.00 1.00 0.50 1.00 1.00 0.43 1.00 0.50 1.00 1.00 0.50 1.00 1.00 1.00 0.33 0.00 1.00 1.00 1.00 1.00 0.33 0.00 1.00 1.00 1.00 1.00 0.33 0.00 1.00	CHR RECE CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_015410 NC_006371 NC_009512 NC_009512 NC_010322 NC_017530 NC_017530 NC_017530 NC_017530 NC_017530 NC_018220 NC_00892 NC_016147 NC_009434 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_017532 NC_010507 NC_010507 NC_010504 NC_012997 NC_013889 NC_011901 NC_012997 NC_013891 NC_016010 NC_003902 NC_007508 NC_007508 NC_017271 NC_002488 NC_010513 NC_010517 NC_010513 NC_0105612 NC_006834	 P. mendocina P. protegens P. putida P. sutata P. spadix P. stutzeri S. suonensis P. syringae P. syringae P. syringae P. syringae P. syringae S. agarivorans S. degradans S. enterica S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. maltophilia S. turnerae X. atonopodis X. campestris X. fastidiosa 	yproteobacteria	0.50 0.04 0.47 0.50 0.64 1.00 0.64 1.00 0.64 0.27 1.00 0.55 0.50 0.50 0.50 0.50 0.60 0.64 1.00 0.47 0.00 0.64 1.00 0.47 0.55 0.50 0.00 0.64 1.00 0.47 0.55 0.50 0.60 0.64 1.00 0.55 0.50 0.50 0.50 0.64 1.00 0.55 0.50 0.50 0.50 0.55 0.50 0.64 1.00 0.45 0.55 0.50 0.50 0.64 1.00 0.47 0.55 0.50 0.50 0.64 1.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.00 0.47 0.55 0.50 0.55 0.40 0.40 0.44 0.44 0.67 0.55 0.40 0.40 0.44 0.55 0.55 0.50 0.55 0.50 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.47 0.55 0.55 0.40 0.40 0.47 0.55 0.40 0.40 0.47 0.55 0.40 0.40 0.47 0.55 0.40 0.55 0.33 0.40 0.55 0.33 0.40 0.55 0.55 0.33 0.40 0.55 0.55 0.55 0.40 0.40 0.40 0.55 0.55 0.55 0.40 0.40 0.40 0.55 0.55 0.55 0.55 0.55 0.55 0.40 0.40 0.55 0

CHR	NC_007705	X. oryzae	γ proteobacteria	0.18	CHR	NC_017308	$C. \ pseudotuberculos is$	Actinobacteridae	1.00
CHR	NC_010717	X. oryzae	γ proteobacteria	0.55	CHR	NC_017462	$C. \ pseudotuberculos is$	Actinobacteridae	1.00
CHR	NC_017267	X. oryzae	γ proteobacteria	0.18	CHR	NC_017730	C. pseudotuberculosis	Actinobacteridae	1.00
PL	NC_017074	S. ruminantium	Negativicutes	0.95	CHR	NC_017945	C. pseudotuberculosis	Actinobacteridae	1.00
almatan	. 4 at a b : 1:4				CHR	NC_015672	C. pseudotuberculosis	Actinobacteridae	1.00
CUD	: 4 stability in	A familiaria	A siding in shids a	0.72	CHR	NC_015692	C. resistens	Actinobacteridae	1.00
CHR	NC 014550	A. jerrooxidans	Actinobacteridae	1.00	CHR	NC 017317	C. ulcerans	Actinobacteridae	1.00
CHB	NC 008711	A aurescens	Actinobacteridae	1.00	CHR	NC_018101	C. ulcerans	Actinobacteridae	1.00
PL	NC 008713	A. aurescens	Actinobacteridae	0.50	CHR	NC_010545	C. urealuticum	Actinobacteridae	1.00
CHR	NC_008578	A. celluloluticus	Actinobacteridae	1.00	CHR	NC_015859	C. variabile	Actinobacteridae	1.00
CHR	NC_011886	A. chlorophenolicus	Actinobacteridae	1.00	CHR	NC_008278	F. alni	Actinobacteridae	1.00
CHR	NC_014218	A. haemolyticum	Actinobacteridae	1.00	CHR	NC_007777	Frankia sp.	Actinobacteridae	1.00
\mathbf{PL}	NC_010852	A. mediterranei	Actinobacteridae	1.00	CHR	NC_009921	Frankia sp.	Actinobacteridae	1.00
CHR	NC_014318	A. mediterranei	Actinobacteridae	1.00	CHR	NC_014666	Frankia sp.	Actinobacteridae	1.00
CHR	NC_017186	A. mediterranei	Actinobacteridae	1.00	CHR	NC_015656	Frankia symbiont	Actinobacteridae	1.00
CHR	NC_018266	A. mediterranei	Actinobacteridae	1.00	CHR	NC_013441	G. bronchialis	Actinobacteridae	1.00
CHR	NC_013093	A. mirum	Actinobacteridae	1.00	PL	NC_013442	G. bronchialis	Actinobacteridae	1.00
CHR	NC_017093	A. missouriensis	Actinobacteridae	1.00	CHR	NC_013757	G. obscurus	Actinobacteridae	1.00
DI	NC_015145	A. phenanthrenivorans	Actinobacteridae	1.00	CHR	NC 018581	G. polyisoprenivorans	Actinobacteridae	1.00
CHB	NC 017802	A. phenanthrenivoruns	Actinobacteridae	1.00	CHR	NC 013721	Goraonia sp.	Actinobacteridae	1.00
PL.	NC 008539	Arthrohacter SD	Actinobacteridae	0.57	CHB	NC 014644	G. vaginalis	Actinobacteridae	1.00
CHB	NC 008541	Arthrobacter sp.	Actinobacteridae	1.00	CHR	NC_017456	G. vaginalis	Actinobacteridae	1.00
CHR	NC_018531	Arthrobacter sp.	Actinobacteridae	1.00	CHR	NC_014830	I. calvum	Actinobacteridae	1.00
CHR	NC_015564	A. subflavus	Actinobacteridae	1.00	\mathbf{PL}	NC_014957	I. pallida	Planctomycetacia	0.90
CHR	NC_008618	B. adolescentis	Actinobacteridae	1.00	CHR	NC_015588	I. variabilis	Actinobacteridae	1.00
CHR	NC_011835	B. animalis	Actinobacteridae	1.00	CHR	NC_013174	J. denitrificans	Actinobacteridae	1.00
CHR	NC_012814	B. animalis	Actinobacteridae	1.00	CHR	NC_013729	K. flavida	Actinobacteridae	1.00
CHR	NC_012815	B. animalis	Actinobacteridae	1.00	CHR	NC_009664	K. radiotolerans	Actinobacteridae	1.00
CHR	NC_017214	B. animalis	Actinobacteridae	1.00	CHR	NC_010617	K. rhizophila	Actinobacteridae	1.00
CHR	NC_017215	B. animalis	Actinobacteridae	1.00	CHR	NC_013169	K. sedentarius	Actinobacteridae	1.00
CHR	NC_017216	B. animalis	Actinobacteridae	1.00	CHR	NC_016109	K. setae	Actinobacteridae	1.00
CHR	NC_017217	B. animalis P. animalis	Actinobacteridae	1.00	CHR	NC 014391	L. xyli M. aurantiaca	Actinobacteridae	1.00
CHR	NC 017866	B. animalis B. animalis	Actinobacteridae	1.00	PL	NC 006910	M rosaria	Actinobacteridae	1.00
CHR	NC_017867	B. animalis	Actinobacteridae	1.00	CHR	NC_014815	Micromonospora sp.	Actinobacteridae	1.00
CHR	NC_018720	B. asteroides	Actinobacteridae	1.00	CHR	NC_010397	M. abscessus	Actinobacteridae	1.00
CHR	NC_014616	B. bifidum	Actinobacteridae	1.00	CHR	NC_015758	M. africanum	Actinobacteridae	1.00
CHR	NC_014638	B. bifidum	Actinobacteridae	1.00	CHR	NC_002944	M. avium	Actinobacteridae	1.00
CHR	NC_017999	B. bifidum	Actinobacteridae	1.00	CHR	NC_008595	M. avium	Actinobacteridae	1.00
CHR	NC_017218	B. breve	Actinobacteridae	1.00	CHR	NC_002945	M. bovis	Actinobacteridae	nan
CHR	NC_012669	B. cavernae	Actinobacteridae	1.00	CHR	NC_008769	M. bovis	Actinobacteridae	1.00
CHR	NC_013714	B. dentium	Actinobacteridae	1.00	CHR	NC_012207	M. bovis	Actinobacteridae	1.00
CHR	NC 004207	B. Jaecium	Actinobacteridae	1.00	CHR	NC 015848	M. consttii	Actinobacteridae	1.00
CHR	NC 010816	B. longum	Actinobacteridae	1.00	CHR	NC 018027	M. chubuense	Actinobacteridae	1.00
CHR	NC_011593	B. longum	Actinobacteridae	1.00	CHR	NC_014246	M. curtisii	Actinobacteridae	1.00
CHR	NC_014169	B. longum	Actinobacteridae	1.00	CHR	NC_009338	M. gilvum	Actinobacteridae	1.00
CHR	NC_014656	B. longum	Actinobacteridae	1.00	\mathbf{PL}	NC_014811	M. gilvum	Actinobacteridae	0.05
CHR	NC_{015052}	B. longum	Actinobacteridae	1.00	CHR	NC_014814	M. gilvum	Actinobacteridae	1.00
CHR	NC_015067	B. longum	Actinobacteridae	1.00	CHR	NC_018612	M. indicus	Actinobacteridae	1.00
CHR	NC_017219	B. longum	Actinobacteridae	1.00	CHR	NC_016946	M. intracellulare	Actinobacteridae	1.00
CHR	NC_017221	B. longum	Actinobacteridae	1.00	CHR	NC_016947	M. intracellulare	Actinobacteridae	1.00
PL	NC_003527	B. pseudocatenulatum	Actinobacteridae	1.00	CHR	NC_016948	M. intracellulare	Actinobacteridae	1.00
CHR	NC_010943	B. saxoosiaens C. aaidimbila	Actinobacteridae	1.00	CHR	NC 002509	M. lenrae	Actinobacteridae	1.00
CHR	NC 012590	C. aurimucosum	Actinobacteridae	1.00	CHB	NC 011896	M. leprae	Actinobacteridae	1.00
CHR	NC 002935	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_012803	M. luteus	Actinobacteridae	1.00
CHR	NC_016782	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_010612	M. marinum	Actinobacteridae	1.00
CHR	NC_016783	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_017955	M. marinus	Actinobacteridae	1.00
CHR	NC_016785	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_018150	M. massiliense	Actinobacteridae	1.00
CHR	NC_016786	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_015635	M. phosphovorus	Actinobacteridae	1.00
CHR	NC_016787	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_016604	M. rhodesiae	Actinobacteridae	1.00
CHR	NC_016788	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_008596	M. smegmatis	Actinobacteridae	1.00
CHR	NC_016789	C. diphtheriae	Actinobacteridae	1.00	CHR	NC 015125	M testacour	Actinobacteridae	1.00
Сир	NC 016790	C. diphtheriae	Actinobacteridae	1.00	CHR	NC 000062	M tuberculosis	Actinobacteridae	1.00
CHR	NC 016800	C. diphtheriae	Actinobacteridae	1.00	CHB	NC 002755	M. tuberculosis	Actinobacteridae	1.00
CHB	NC 016801	C diphtheriae	Actinobacteridae	1.00	CHR	NC 009525	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_016802	C. diphtheriae	Actinobacteridae	1.00	CHR	NC_009565	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_004369	C. efficiens	Actinobacteridae	1.00	CHR	NC_012943	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_015514	C. fimi	Actinobacteridae	1.00	CHR	NC_016768	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_014151	C. flavigena	Actinobacteridae	1.00	CHR	NC_016934	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_015671	C. gilvus	Actinobacteridae	1.00	CHR	NC_017026	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_003450	C. glutamicum	Actinobacteridae	1.00	CHR	NC_017522	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_006958	C. glutamicum	Actinobacteridae	nan	CHR	NC_017523	M. tuberculosis	Actinobacteridae	1.00
CHR DT	NC 000242	C. giutamicum	Actinobacteridae	1.00	Снк	NC 017524	M. tuberculosis	Actinobacteridae	1.00
снв	NC 007164	0. gaaamicam C ieikeinm	Actinobacteridae	1.00	CHR	NC 018078	M tuberculosis	Actinobacteridae	1.00
CHR	NC.012704	C. kroppenstedtii	Actinobacteridae	1.00	CHR	NC_018143	M. tuberculosis	Actinobacteridae	1.00
CHR	NC_009480	C. michiganensis	Actinobacteridae	1.00	CHR	NC_008611	M. ulcerans	Actinobacteridae	1.00
CHR	NC_010407	C. michiganensis	Actinobacteridae	1.00	CHR	NC_008726	M. vanbaalenii	Actinobacteridae	1.00
CHR	NC_014329	C. pseudotuberculosis	Actinobacteridae	1.00	CHR	NC_008146	Mycobacterium sp.	Actinobacteridae	1.00
CHR	NC_016781	$C.\ pseudotuberculos is$	Actinobacteridae	1.00	CHR	NC_008705	Mycobacterium sp.	Actinobacteridae	1.00
CHR	NC_016932	$C.\ pseudotuberculos is$	Actinobacteridae	nan	CHR	NC_009077	Mycobacterium sp.	Actinobacteridae	1.00
CHR	NC_017031	$C. \ pseudotuberculosis$	Actinobacteridae	1.00	CHR	NC_015576	Mycobacterium sp.	Actinobacteridae	1.00
CHR	NC_017300	C. pseudotuberculosis	Actinobacteridae	1.00	CHR	NC_017904	Mycobacterium sp.	Actinobacteridae	1.00
CHR	NC_017301	C. pseudotuberculosis	Actinobacteridae	1.00	CHR	NC_018524	N. alba N. brasilismois	Actinobacteridae	1.00
СНР	NC 017205	C. pseudotuberculosis	Acunobacteridae	1.00	CHP	NC 016887	N. curiaciaeoraica	Actinobacteridae	1.00
CHR	NC 017306	C. pseudotuberculosis	Actinobacteridae	1.00	CHR	NC_014210	N. dassonvillei	Actinobacteridae	1.00
CHR	NC_017307	C. pseudotuberculosis	Actinobacteridae	1.00	RECE	NC_014211	N. dassonvillei	Actinobacteridae	1.00

CHR	$NC_{-}006361$	N. farcinica	Actinobacteridae	1.00	CHR	NC_018643	α proteobacteriaacterium	α proteobacteria	1.00
CHR	NC_013235	N. multipartita	Actinobacteridae	nan	CHR	NC_018644	α proteobacteriaacterium	α proteobacteria	1.00
PL	NC_008697	Nocardioides sp.	Actinobacteridae	0.53	CHR	NC_009445	Bradyrhizobium sp.	α proteobacteria	1.00
CHR	NC 019395	P acidinronionici	Actinobacteridae	1.00	CHR	NC 017082	Bradyrhizobium sp. Bradyrhizobium sp.	aproteobacteria	1.00
CHR	NC_006085	P. acnes	Actinobacteridae	1.00	CHR	NC_006932	B. abortus	αproteobacteria	1.00
CHR	NC_014039	P. acnes	Actinobacteridae	1.00	CHR	NC_010742	B. abortus	α proteobacteria	1.00
CHR	NC_016511	P. acnes	Actinobacteridae	1.00	CHR	NC_016795	B. abortus	α proteobacteria	1.00
CHR	NC_016512	P. acnes	Actinobacteridae	1.00	CHR	NC_008783	B. bacilliformis	α proteobacteria	1.00
CHR	NC_016516 NC_017534	P. acnes P. acnes	Actinobacteridae	1.00	CHR	NC_010103 NC_016778	B. canis B. canis	αproteobacteria	1.00
CHR	NC_017535	P. acnes	Actinobacteridae	1.00	CHR	NC_014932	B. clarridgeiae	αproteobacteria	1.00
CHR	NC_017550	P. acnes	Actinobacteridae	1.00	CHR	NC_012846	B. grahamii	α proteobacteria	1.00
CHR	NC_018707	P. acnes	Actinobacteridae	1.00	CHR	$NC_{-}005956$	B. henselae	α proteobacteria	1.00
CHR	NC_015312	P. dioxanivorans	Actinobacteridae	1.00	CHR	NC_010581	B. indica	α proteobacteria	1.00
CHR	NC_014215 NC_018142	P. freudenreichii P. propionieum	Actinobacteridae	1.00	CHR	NC_004463	B. japonicum B. japonicum	αproteobacteria	1.00
CHR	NC_014643	R. dentocariosa	Actinobacteridae	1.00	CHR	NC_003317	B. melitensis	αproteobacteria	1.00
CHR	NC_014659	R. equi	Actinobacteridae	1.00	CHR	NC_007618	B. melitensis	αproteobacteria	1.00
PL	NC_005073	R. erythropolis	Actinobacteridae	0.14	CHR	NC_012441	B. melitensis	α proteobacteria	1.00
PL	NC_007491	R. erythropolis	Actinobacteridae	0.71	CHR	NC_017244	B. melitensis	α proteobacteria	1.00
CHR	NC_012490	R. erythropolis P. iostii	Actinobacteridae	1.00	CHR	NC_017246	B. melitensis	αproteobacteria	1.00
CHR	NC 013715	R. jostii R. mucilaainosa	Actinobacteridae	1.00	CHR	NC 013119	B. metitensis B. microti	aproteobacteria	1.00
CHR	NC_012522	R. opacus	Actinobacteridae	1.00	CHR	NC_009505	B. ovis	αproteobacteria	1.00
CHR	NC_010168	R. salmoninarum	Actinobacteridae	1.00	CHR	NC_015857	B. pinnipedialis	α proteobacteria	1.00
PL	$NC_{-}006571$	S. albulus	Actinobacteridae	1.00	CHR	$NC_{-}005955$	B. quintana	α proteobacteria	1.00
CHR	NC_009953	S. arenicola	Actinobacteridae	1.00	CHR	NC_018533	B. quintana	α proteobacteria	1.00
CHR	NC_003155 NC_016582	S. avermitilis S. hinachengaensis	Actinobacteridae	1.00	CHR	NC_014375 NC_004310	B. subvibrioides	αproteobacteria	1.00
CHR	NC_016111	S. cattleya	Actinobacteridae	1.00	CHR	NC_004310	B. suis	αproteobacteria	1.00
PL	NC_016113	S. cattleya	Actinobacteridae	1.00	CHR	NC_016797	B. suis	α proteobacteria	1.00
PL	NC_{017585}	S. cattleya	Actinobacteridae	1.00	CHR	NC_{017251}	B. suis	α proteobacteria	1.00
CHR	NC_017586	S. cattleya	Actinobacteridae	1.00	CHR	NC_010161	B. tribocorum	α proteobacteria	1.00
PL DI	NC_007392	S. citri	Mollicutes	1.00	CHR	NC_002696	C. crescentus	αproteobacteria	1.00
CHR	NC_003888	S. coelicolor	Actinobacteridae	1.00	CHR	NC_010338	Caulobacter sp.	aproteobacteria	1.00
PL	NC_002112	S. cyaneus	Actinobacteridae	1.00	CHR	NC_008254	Chelativorans sp.	αproteobacteria	1.00
CHR	NC_009142	S. erythraea	Actinobacteridae	1.00	CHR	NC_012985	Cand. Liberibacter	α proteobacteria	1.00
CHR	NC_016114	S. flavogriseus	Actinobacteridae	1.00	CHR	NC_014774	Cand. Liberibacter	α proteobacteria	1.00
PL	NZ_CM001485	S. glauca	Actinobacteridae	1.00	CHR	NC_015722	Cand. Midichloria	αproteobacteria	1.00
CHR	NC 017765	S. griseus S. huarosconicus	Actinobacteridae	1.00	CHR	NC 015380	Cand. Pelagibacter	aproteobacteria	1.00
CHR	NC_013521	S. keddieii	Actinobacteridae	1.00	CHR	NC_014010	Cand. Puniceispirillum	αproteobacteria	1.00
PL	NC_006400	S. kunkelii	Mollicutes	1.00	CHR	NC_014100	C. segnis	α proteobacteria	1.00
CHR	NC_013947	$S.\ nassauens is$	Actinobacteridae	1.00	CHR	NC_009952	D. shibae	α proteobacteria	1.00
CHR	NC_013595	S. roseum	Actinobacteridae	1.00	CHR	NC_007354	E. canis	α proteobacteria	1.00
CHR	NC_014168 NC_013929	S. rotundus S. scabiei	Actinobacteridae	1.00	CHR	NC 007799	E. chaffeensis E. litoralis	αproteobacteria	1.00
PL	NC_006911	Streptomyces sp.	Actinobacteridae	0.14	CHR	NC_005295	E. ruminantium	αproteobacteria	1.00
CHR	NC_015953	Streptomyces sp.	Actinobacteridae	1.00	CHR	NC_006831	E. ruminantium	α proteobacteria	1.00
CHR	NC_009380	S. tropica	Actinobacteridae	1.00	CHR	$NC_{-}006832$	$E. \ ruminantium$	α proteobacteria	nan
CHR	NC_018750	S. venezuelae	Actinobacteridae	1.00	CHR	NC_008343	G. bethesdensis	α proteobacteria	1.00
CHR	NC_015957	S. violaceusniger	Actinobacteridae	1.00	CHR	NC_010125	G. diazotrophicus	αproteobacteria	1.00
CHR	NC 014165	T hispora	Actinobacteridae	1.00	CHR	NC 006677	G. anazotrophicus G. orudans	aproteobacteria	1.00
CHR	NC_013510	T. curvata	Actinobacteridae	1.00	CHR	NC_019396	G. oxydans	αproteobacteria	1.00
CHR	NC_007333	T. fusca	Actinobacteridae	1.00	CHR	$NC_{-}016027$	G. xylinus	α proteobacteria	1.00
CHR	NC_014158	T. paurometabola	Actinobacteridae	1.00	CHR	NC_015717	Hyphomicrobium sp.	α proteobacteria	1.00
CHR	NC_004551	T. whipplei T. whipplei	Actinobacteridae	1.00	CHR	NC_012982	H. baltica	αproteobacteria	1.00
CHR	NC 015434	1. wnippiei V maris	Actinobacteridae	1.00	CHR	NC 008358	H. aenitripicans H. nentunium	αproteobacteria	1.00
CHR	NC_013530	X. cellulosilytica	Actinobacteridae	1.00	CHR	NC_007802	Jannaschia sp.	αproteobacteria	1.00
CHR	NC_013739	C. woesei	Rubrobacridae	1.00	CHR	$NC_{-}014625$	K. vulgare	α proteobacteria	1.00
					CHR	NC_017384	K. vulgare	α proteobacteria	1.00
cluster :	5 stability mea	sure : 0.960160		1.00	CHR	NC_010511	Methylobacterium sp.	αproteobacteria	1.00
CHR	NC 013209	A. pasteurianus	aproteobacteria	1.00	CHR	NC 016026	Meingiocysiis sp. M. aeruginosavorus	αproteobacteria	1.00
CHR	NC_017108	A. pasteurianus	αproteobacteria	1.00	CHR	NC_011757	M. chloromethanicum	αproteobacteria	1.00
CHR	NC_017111	A. pasteurianus	α proteobacteria	1.00	CHR	NC_014923	M. ciceri	α proteobacteria	1.00
CHR	NC_017121	$A. \ pasteurianus$	α proteobacteria	1.00	CHR	NC_010172	M. extorquens	α proteobacteria	1.00
CHR	NC_017125	A. pasteurianus	α proteobacteria	1.00	CHR	NC_012808	M. extorquens	α proteobacteria	1.00
CHR	NC_017146	A. pasteurianus	αproteobacteria	1.00	PL	NC_012809	M. extorquens	αproteobacteria	nan 0.50
CHR	NC_009484	A. cruptum	aproteobacteria	1.00	CHR	NC_012988	M. extorquens	αproteobacteria	1.00
CHR	NC_015186	A. multivorum	α proteobacteria	1.00	CHR	NC_002678	M. loti	α proteobacteria	1.00
CHR	NC_{013532}	A. centrale	α proteobacteria	1.00	CHR	$NC_{-}007626$	M. magneticum	α proteobacteria	1.00
CHR	NC_004842	A. marginale	α proteobacteria	1.00	CHR	NC_008347	M. maris	α proteobacteria	1.00
CHR	NC_012026	A. marginale	α proteobacteria	1.00	CHR	NC_015675	M. nodulans	αproteobacteria	1.00
CHR	NC 009937	A. pnayocytophilum	αproteopacteria	1.00	CHR	NC_010725	M. opportunistum M. populi	αproteobacteria	1.00
CHR	NC_003062	A. fabrum	αproteobacteria	1.00	CHR	NC_010505	M. radiotolerans	α proteobacteria	1.00
CHR	NC_011985	$A.\ radiobacter$	α proteobacteria	1.00	CHR	NC_011666	M. silvestris	α proteobacteria	1.00
CHR	NC_011989	A. vitis	α proteobacteria	1.00	CHR	NC_015580	Novosphingobium sp.	α proteobacteria	1.00
CHR	NC_015183	Agrobacterium sp.	α proteobacteria	1.00	CHR	NC_007794	N. aromaticivorans	α proteobacteria	1.00
CHR	NC 014816	A. excentricus	αproteobacteria	1.00	CHR	NC 007406	N. winoaradskui	aproteobacteria	1 nan
PL	NC_016594	A. brasilense	αproteobacteria	1.00	CHR	NC_009667	O. anthropi	aproteobacteria	1.00
CHR	NC_016617	A. brasilense	α proteobacteria	1.00	\mathbf{CHR}	NC_011386	O. carboxidovorans	α proteobacteria	1.00
PL	NC_016587	A. lipoferum	α proteobacteria	1.00	\mathbf{CHR}	NC_015684	$O.\ carboxidovorans$	α proteobacteria	1.00
CHR	NC_016622	A. lipoferum	α proteobacteria	1.00	CHR	NC_017538	O. carboxidovorans	α proteobacteria	1.00
CHR	NC_013854	Azospirillum sp.	α proteobacteria	1.00	CHR	NC_016642	Pseudovibrio sp.	α proteobacteria	1.00
PL	NC_013858	Azospirillum sp.	α proteobacteria	1.00	OHR	INC_014414	1. vermuaensis	αproteopacteria	1 1.00

253

CHR	NC_008686	P. denitrificans	αproteobacteria	1.00	CHR	NC_010981	Wolbachia endosymbiont	αproteobacteria	1.00
RECE	NC_008687	P. denitrificans	α proteobacteria	1.00	CHR	NC_018267	Wolbachia endosymbiont	α proteobacteria	1.00
CHR	NC_018286	P. gallaeciensis	α proteobacteria	1.00	\mathbf{PL}	NC_003425	W. glossinidia	α proteobacteria	1.00
CHR	NC_018290	P. gallaeciensis	α proteobacteria	1.00	CHR	NC_009720	X. autotrophicus	α proteobacteria	1.00
CHR	NC_015259	P. gilvum	α proteobacteria	1.00	CHR	NC_006526	Z. mobilis	α proteobacteria	1.00
CHR	NC_016078	P. halotolerans	α proteobacteria	1.00	CHR	NC_013355	Z. mobilis	α proteobacteria	1.00
CHR	NC_009719	$P.\ lavamentivorans$	α proteobacteria	1.00	CHR	NC_015709	Z. mobilis	α proteobacteria	1.00
CHR	NC_011144	P. zucineum	α proteobacteria	1.00	CHR	NC_017262	Z. mobilis	α proteobacteria	1.00
CHR	NC_008044	Ruegeria sp.	α proteobacteria	1.00	CHR	NC_018145	Z. mobilis	α proteobacteria	1.00
CHR	NC_012633	R. africae	α proteobacteria	0.93					
CHR	NC_009881	R. akari	α proteobacteria	0.67	cluster :	6 stability me	asure : 0.962091		
CHR	NC_017058	R. australis	α proteobacteria	0.91	CHR	NC_018011	A. finegoldii	Bacteroidia	1.00
CHR	NC_007940	R. bellii	α proteobacteria	1.00	CHR	NC_003228	B. fragilis	Bacteroidia	1.00
CHR	NC_009883	R. bellii	αproteobacteria	1.00	CHR	NC_006347	B. fragilis	Bacteroidia	1.00
CHR	NC_009879	R. canadensis	αproteobacteria	1.00	CHR	NC-016776	B. fragilis	Bacteroidia	1.00
CHR	NC 014024	R. canadensis	aproteobacteria	1.00	CHR	NC_014933	B. nelcogenes	Bacteroidia	1.00
CHR	NC 0114034	R. centenum	aproteobacteria	1.00	CHR	NC 004662	B. saturnitions	Bacteroidia	1.00
CHB	NC 003103	R conorii	aproteobacteria	0.86	DI	NC 004702	B. thetaiotaomicron	Bacteroidia	0.52
CHB	NC 008209	R denitrificans	aproteobacteria	1.00	CHB	NC 009614	B. milaatus	Bacteroidia	1.00
CHR	NC 007761	R. etli	αproteobacteria	1.00	CHR	NC 011565	Cand Azobacteroides	Bacteroidia	1.00
CHR	NC_010994	R. etli	α proteobacteria	1.00	CHR	NC_015160	O. splanchnicus	Bacteroidia	1.00
CHR	NC_007109	R. felis	α proteobacteria	0.93	CHR	NC_015501	P. asaccharolytica	Bacteroidia	1.00
CHR	NC_015866	R. heilongjiangensis	α proteobacteria	0.93	CHR	NC_015311	P. denticola	Bacteroidia	1.00
CHR	NC_016050	R. japonica	α proteobacteria	0.91	CHR	NC_009615	P. distasonis	Bacteroidia	1.00
CHR	NC_008380	R. leguminosarum	α proteobacteria	nan	CHR	NC_002950	P. gingivalis	Bacteroidia	1.00
CHR	NC_011369	$R. \ leguminos arum$	α proteobacteria	1.00	CHR	NC_010729	P. gingivalis	Bacteroidia	1.00
CHR	$NC_{-}012850$	R. leguminosarum	α proteobacteria	1.00	CHR	NC_015571	P. gingivalis	Bacteroidia	1.00
CHR	NC_015730	R. litoralis	α proteobacteria	1.00	\mathbf{CHR}	NC_017860	P. intermedia	Bacteroidia	1.00
CHR	NC_009900	R. massiliae	α proteobacteria	1.00	RECE	NC_017861	P. intermedia	Bacteroidia	1.00
CHR	NC_016931	R. massiliae	α proteobacteria	1.00	CHR	NC_014370	P. melaninogenica	Bacteroidia	1.00
CHR	NC_017043	R. montanensis	α proteobacteria	1.00	RECE	NC_014371	P. melaninogenica	Bacteroidia	1.00
CHR	NC_005296	R. palustris	αproteobacteria	1.00	CHR	NC_014734	P. propionicigenes	Bacteroidia	1.00
CHP	NC 007005	n. parastris	aproteobacteria	1.00	CHR	NC_014033	r. ruminicola	Bacteroidia	1.00
CHR	NC 007925	R. palustris	αproteobacteria	1.00	CHR	NC_016610	T. forsythia	Chlanchia	1.00
CHR	NC 008425	R. palustris	aproteobacteria	1.00	CHR	NC_011020	C. indiassium	Castanhania	1.00
CHR	NC 011004	R palustris	aproteobacteria	1.00	CHR	NC 008255	B. butchinsonii	Cytophagia	1.00
CHB	NC 014834	R nalustris	aproteobacteria	1.00	CHR	NC 015914	C. marinum	Cytophagia	1.00
CHR	NC 017044	R. parkeri	αproteobacteria	1.00	CHB	NC 013037	D fermentans	Cytophagia	1.00
CHR	NC_012730	B. peacockii	αproteobacteria	0.93	CHB	NC 018748	E oligotrophica	Cytophagia	1.00
CHR	NC_016930	R. philipii	α proteobacteria	0.91	CHR	NC_018018	F. litoralis	Cytophagia	1.00
CHR	NC_017059	R. photometricum	α proteobacteria	1.00	CHR	NC_014655	L. byssophila	Cytophagia	1.00
CHR	NC_003911	R. pomeroyi	α proteobacteria	1.00	CHR	NC_014759	M. tractuosa	Cytophagia	1.00
CHR	NC_000963	R. prowazekii	α proteobacteria	1.00	CHR	NC_015703	R. slithyformis	Cytophagia	1.00
CHR	NC_017048	R. prowazekii	α proteobacteria	0.91	\mathbf{PL}	NC_015704	R. slithyformis	Cytophagia	1.00
CHR	NC_017049	R. prowazekii	α proteobacteria	0.93	CHR	NC_013730	S. linguale	Cytophagia	1.00
CHR	NC_017050	R. prowazekii	α proteobacteria	0.86	CHR	NC_018013	A. sublithincola	Flavobacteriia	1.00
CHR	NC_017051	R. prowazekii	α proteobacteria	1.00	CHR	NC_013418	Blattabacterium sp.	Flavobacteriia	1.00
CHR	NC_017056	R. prowazekii	α proteobacteria	0.91	CHR	NC_013454	Blattabacterium sp.	Flavobacteriia	1.00
CHR	NC_017057	R. prowazekii	α proteobacteria	1.00	CHR	NC_016146	Blattabacterium sp.	Flavobacteriia	1.00
CHR	NC_017560	R. prowazeku	αproteobacteria	0.91	CHR	NC_016621	Blattabacterium sp.	Flavobacteriia	1.00
CHR	NC_017042	R. rnipicepnali	αproteobacteria	0.91	CHR	NC_017924	Blattabacterium sp.	Flavobacteriia	1.00
CHR	NC 010262	R. rickettsii	aproteobacteria	1.00	CHR	NC_014934	C. algicola	Flavobacteriia	1.00
CHR	NC 016908	R. rickettsii	aproteobacteria	0.91	CHR	NC 015846	C. amimonaua	Flavobacteriia	1.00
CHB	NC 016909	R rickettsii	aproteobacteria	0.55	CHR	NC 015167	C. lutica	Flavobacteriia	1.00
CHR	NC_016911	R. rickettsii	αproteobacteria	0.86	CHB	NC 013162	C. ochracea	Flavobacterija	1.00
CHR	NC_016913	R. rickettsii	αproteobacteria	0.86	CHR	NC 010118	C. Sulcia	Flavobacterija	1.00
CHR	NC_016914	R. rickettsii	α proteobacteria	0.93	CHR	NC_013123	C. Sulcia	Flavobacteriia	1.00
CHR	NC_016915	R. rickettsii	α proteobacteria	0.93	CHR	NC_014004	C. Sulcia	Flavobacteriia	1.00
CHR	$NC_{-}007643$	R. rubrum	α proteobacteria	1.00	CHR	NC_014499	C. Sulcia	Flavobacteriia	1.00
CHR	NC_017584	R. rubrum	α proteobacteria	1.00	CHR	NC_013062	Flavobacteriaceae sp.	Flavobacteriia	1.00
CHR	NC_016639	R. slovaca	α proteobacteria	0.91	CHR	NC_016001	$F.\ branchiophilum$	Flavobacteriia	1.00
CHR	NC_017065	R. slovaca	α proteobacteria	1.00	CHR	NC_016510	F. columnare	Flavobacteriia	1.00
CHR	NC_007493	R. sphaeroides	α proteobacteria	1.00	CHR	NC_017025	F. indicum	Flavobacteriia	1.00
CHR	NC-009049	R. sphaeroides	α proteobacteria	1.00	CHR	NC_009441	F. johnsoniae	F lavobacteriia	1.00
CHR	NC 011062	n. spnaeroides	aproteobacteria	1.00	CHR	NC_009613	F. psychrophilum	r lavobacteriia	1.00
CHR	NC 006142	R. sphaeronaes	aproteobacteria	1.00	CHR	NC_015321	F. taffensis	Flavobacteriia	1.00
CHP	NC 017062	R tunhi	aproteobacteria	0.81	СНР	NC 015406	G. juiselli Krakinghaster sp	Flavobacteriia	1.00
CHB	NC-017066	R. tuphi	aproteobacteria	0.91	CHP	NC 015638	Lacinutrix sp.	Flavobacterija	1.00
CHR	NC 014664	R. vannielii	αproteobacteria	1.00	CHB	NC 014472	Maribacter sp.	Flavobacterija	1.00
CHR	NC_015976	Sphingobium sp.	αproteobacteria	1.00	CHR	NC_015945	M. ruestringensis	Flavobacterija	1.00
CHR	NC_008048	S. alaskensis	α proteobacteria	1.00	CHR	NC_016599	O. hongkongensis	Flavobacteriia	1.00
CHR	NC_015593	S. chlorophenolicum	α proteobacteria	1.00	CHR	NC_018016	O. rhinotracheale	Flavobacteriia	1.00
RECE	NC_015594	S. chlorophenolicum	α proteobacteria	0.41	CHR	NC_018721	P. torquis	Flavobacteriia	1.00
\mathbf{CHR}	NC_012587	S. fredii	α proteobacteria	1.00	CHR	NC_014738	R. anatipestifer	Flavobacteriia	1.00
CHR	NC_016812	S. fredii	α proteobacteria	1.00	CHR	NC_017045	R. an a tipe stifer	Flavobacteriia	1.00
CHR	NC_018000	S. fredii	α proteobacteria	1.00	CHR	NC_017569	$R. \ an a tipe stifer$	Flavobacteriia	1.00
CHR	NC_014006	S. japonicum	α proteobacteria	1.00	CHR	NC_018609	R. an a tipe stifer	Flavobacteriia	1.00
CHR	NC_009636	S. medicae	α proteobacteria	1.00	CHR	NC_013222	R. biformata	Flavobacteriia	1.00
CHR	NC_003047	S. meliloti	α proteobacteria	1.00	CHR	NC_015144	W. virosa	Flavobacteriia	1.00
CHR	NC_015590	S. meliloti	α proteobacteria	1.00	CHR	NC_015844	Z. galactanivorans	Flavobacteriia	nan
CHR	NC 017322	5. melliloti	aproteobacteria	1.00	CHR	NC_014041	∠. projunda	r lavobacteriia	1.00
CHP	NC 018700	S. meliloti	aproteobacteria	1.00	СИВ	NC 012122	1. atoum C. ninensis	ignavibacteria Sphingobacteria	1.00
CHB	NC 014217	S. novella	aproteobacteria	1.00	СНР	NC 015510	. privensis H hudrossis	Sphingobacteria	1.00
CHR	NC_009511	S. wittichii	aproteobacteria	1.00	CHR	NC 016609	N koreensis	Sphingobacteria	1.00
CHR	NC_017956	T. mobilis	α proteobacteria	1.00	CHR	NC_013061	P. heparinus	Sphingobacteria	1.00
CHR	NC_012416	Wolbachia sp.	α proteobacteria	1.00	CHR	NC_015177	P. saltans	Sphingobacteria	1.00
CHR	NC_002978	Wolbachia endosymbiont	α proteobacteria	1.00	CHR	NC_017770	S. canadensis	Sphingobacteria	1.00
CHR	NC_006833	Wolbachia endosymbiont	α proteobacteria	1.00	CHR	NC_015277	Sphingobacterium sp.	Sphingobacteria	1.00

CHR	NC_016940	S. grandis	Sphingobacteria	1.00	PL	NC_017647	E. coli	γ proteobacteria	1.00
CHR	NC_010830	Cand. Amoebophilus	Bacteroidetes I.S.	1.00	\mathbf{PL}	$NC_{-}011752$	E. coli	γ proteobacteria	1.00
CHR	NC_018605	Cardinium endosymbiont	Bacteroidetes I.S.	1.00	PL	NC_019000	E. coli	γ proteobacteria	1.00
<u> </u>					PL	NC_013942	E. coli	γ proteobacteria	1.00
cluster	: 7 stability me	easure : 0.960840	A 1	0.01	PL	NC_007635	E. coli	γ proteobacteria	1.00
	NC_004963	M. celatum	Actinobacteridae	0.91	PL	NC_012728	E. coli E. coli	γ proteobacteria	1.00
PL	NC 010374	Methylobacterium sp.	αproteobacteria	nan	PL	NC_013507	E. coli	γproteobacteria	1.00
PL	NC_012109	D. autotrophicum	δ proteobacteria	0.77	\mathbf{PL}	NC_017657	E. coli	γ proteobacteria	1.00
PL	NC_004843	B. aphidicola	γ proteobacteria	1.00	\mathbf{PL}	NC_017653	E. coli	γ proteobacteria	1.00
PL	NC_004464	C. freundii	γ proteobacteria	1.00	PL	NC_014234	E. coli	$\gamma \mathrm{proteobacteria}$	1.00
PL	NC_013717	C. rodentium	γ proteobacteria	1.00	PL	NC_011812	E. coli	γ proteobacteria	1.00
PL	NC_009780	C. sakazakii	γ proteobacteria	1.00	PL	NC_009837	E. coli	γ proteobacteria	1.00
PL DI	NC_013285	C. turicensis	γproteobacteria	1.00	PL	NC 010558	E. coli E. coli	γ proteobacteria	1.00
PL	NC 013973	E amulovora	vproteobacteria	1.00	PL	NC 007941	E. coli	vproteobacteria	1.00
PL	NC_005246	E. amylovora	γ proteobacteria	1.00	PL	NC_010488	E. coli	γ proteobacteria	1.00
PL	NC_015963	E. asburiae	γ proteobacteria	1.00	\mathbf{PL}	NC_013175	E. coli	γ proteobacteria	1.00
PL	NC_014305	E. billingiae	γ proteobacteria	1.00	PL	NC_{017907}	E. coli	$\gamma \mathrm{proteobacteria}$	1.00
PL	NC_014108	E. cloacae	γ proteobacteria	0.12	PL	NC_011754	E. coli	γ proteobacteria	1.00
PL DI	NC_014107	E. cloacae	γ proteobacteria	1.00		NC_011964	E. coli E. coli	γ proteobacteria	1.00
PL	NC 000788	E. cloacae E. coli	γproteobacteria	1.00	PL PL	NC 011980	E. coli E. coli	γproteobacteria	1.00
PL	NC 009787	E. coli	vproteobacteria	1.00	PL	NC_013122	E. coli	vproteobacteria	1.00
PL	NC_009786	E. coli	γ proteobacteria	1.00	PL	NC_016039	E. coli	γ proteobacteria	1.00
PL	NC_019061	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_014477	E. coli	γ proteobacteria	1.00
PL	NC_019063	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_{011747}	E. coli	γ proteobacteria	1.00
PL	NC_013120	E. coli	γ proteobacteria	1.00	PL	NC_011749	E. coli	γ proteobacteria	1.00
PL	NC_017630	E. coli	γ proteobacteria	1.00	PL	NC_007414	E. coli	γ proteobacteria	1.00
PL DI	NC_002128	E. coli E. coli	γproteobacteria	1.00	PL PI	NC_011350	E. coli E. coli	γ proteobacteria	1.00
PL	NC 009790	E. coli	oproteobacteria	1.00	PL	NC 014384	E. coli	vproteobacteria	1.00
PL	NC_019094	E. coli	γ proteobacteria	1.00	PL	NC_013655	E. coli	γ proteobacteria	1.00
PL	NC_019095	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_011743	E. fergusonii	γ proteobacteria	1.00
PL	NC_019097	E. coli	γ proteobacteria	1.00	PL	NC_{009425}	Enterobacter sp.	$\gamma \mathrm{proteobacteria}$	1.00
PL	NC_019090	E. coli	γ proteobacteria	1.00	PL	NC_014725	E. tarda	γ proteobacteria	1.00
PL	NC_019071	E. coli	γ proteobacteria	1.00	PL	NC_010699	E. tasmaniensis	γ proteobacteria	1.00
	NC_017627	E. coli E. coli	γ proteobacteria	1.00	PL	NC 018107	E. tasmaniensis K. orutoca	γproteobacteria	0.12
PL	NZ DS999999	E. coli	vproteobacteria	1.00	PL	NC_017541	K. pneumoniae	γproteobacteria	1.00
PL	NC_013010	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_019390	K. pneumoniae	γ proteobacteria	1.00
PL	NC_011419	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_015154	K. pneumoniae	γ proteobacteria	1.00
PL	NC_011413	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_019165	K. pneumoniae	$\gamma \mathrm{proteobacteria}$	1.00
PL	NC_011416	E. coli	γ proteobacteria	1.00	PL	NC_013950	K. pneumoniae	γ proteobacteria	1.00
PL DI	NC_014385	E. coli E. coli	γ proteobacteria	1.00	PL	NC_011281	K. pneumoniae	γ proteobacteria	1.00
PL	NC 014615	E. coli	vproteobacteria	1.00	PL	NC 019154	K. pneumoniae	vproteobacteria	1.00
PL	NC_019044	E. coli	γ proteobacteria	1.00	PL	NC_010886	K. pneumoniae	γ proteobacteria	1.00
PL	NC_019043	E. coli	γ proteobacteria	1.00	\mathbf{PL}	$NC_{-}009650$	K. pneumoniae	γ proteobacteria	1.00
PL	NC_011514	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_016966	K. pneumoniae	$\gamma { m proteobacteria}$	1.00
PL	NC_017637	E. coli	γ proteobacteria	1.00	PL	NC_010726	K. pneumoniae	γ proteobacteria	0.41
PL	NC_017639	E. coli	γ proteobacteria	1.00	PL	NC_019389	K. pneumoniae	γ proteobacteria	1.00
PL	NC 008460	E. coli	γproteobacteria	1.00	PL	NC 013951	K. pneumoniae	oproteobacteria	1.00
PL	NC_004998	E. coli	γ proteobacteria	1.00	PL	NC_016846	K. pneumoniae	γ proteobacteria	1.00
PL	NC_011603	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_013542	K. pneumoniae	γ proteobacteria	1.00
PL	NC_013366	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_009649	K. pneumoniae	γ proteobacteria	1.00
PL	NC_014383	E. coli	γ proteobacteria	1.00	PL	NC_014016	K. pneumoniae	γ proteobacteria	1.00
PL	NC_009602	E. coli	γ proteobacteria	1.00	PL	NC_019155	K. pneumoniae	γ proteobacteria	1.00
	NC_018659	E. coli	γ proteobacteria	1.00	PL	NC_014312	K. pneumoniae	γ proteobacteria	1.00
PL	NC 010409	E. coli	vproteobacteria	1.00	PL	NC_011641	K. pneumoniae	γproteobacteria	1.00
PL	NC_010719	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_009651	K. pneumoniae	γ proteobacteria	nan
PL	NC_014843	E. coli	γ proteobacteria	1.00	\mathbf{PL}	$NC_{-}006625$	K. pneumoniae	γ proteobacteria	0.67
PL	NC_017724	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_017533	P. ananatis	$\gamma \mathrm{proteobacteria}$	1.00
PL	NC_017722	E. coli	γ proteobacteria	1.00	PL	NC_016817	P. ananatis	γ proteobacteria	0.95
PL DI	NC_007675	E. coli E. coli	γ proteobacteria	1.00	PL	NC_017553 NC_014842	P. ananatis Pantosa sp	γ proteobacteria	0.92
PL DI	NC 006671	E. coli E. coli	γproteobacteria	1.00	PL	NC 014839	Pantoea sp.	oproteobacteria	1.00
PL	NC_013370	E. coli	γproteobacteria	1.00	PL	NC_014563	P. vagans	γ proteobacteria	1.00
PL	NC_005327	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_014561	P. vagans	γ proteobacteria	1.00
PL	NC_019037	E. coli	γ proteobacteria	1.00	PL	NC_{015063}	Rahnella sp.	$\gamma \mathrm{proteobacteria}$	1.00
PL	NC_013121	E. coli	γ proteobacteria	1.00	PL	NC_017577	S. baltica	γ proteobacteria	0.73
PL	NC_017665	E. coli	γ proteobacteria	1.00	PL	NC_010660	S. boydii	γ proteobacteria	1.00
PL	NC 013362	E. coli	γproteobacteria	1.00	PL	NC 007607	S. dogun S. dusenteriae	oproteobacteria	1.00
PL	NC_010862	E. coli	γproteobacteria	1.00	PL	NC_003277	S. enterica	γ proteobacteria	1.00
PL	NC_018995	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_011081	$S.\ enterica$	γ proteobacteria	1.00
PL	$NC_{-}007365$	E. coli	γ proteobacteria	1.00	\mathbf{PL}	$\rm NC_014476$	S.~enterica	γ proteobacteria	1.00
PL	NC_019093	E. coli	γ proteobacteria	1.00	PL	NC_006816	S. enterica	γ proteobacteria	1.00
PL	NC_014232	E. coli E. coli	γ proteobacteria	1.00	PL	NC_011092	S. enterica	γ proteobacteria	1.00
PL	NC-013354	E. coli	γproteobacteria	1.00	PL PI	NC 017718	5. enterica S enterica	γ proteobacteria	1.00
PL	NC_019072	E. coli	vproteobacteria	1.00	PL	NC_019117	S. enterica	γproteobacteria	1.00
PL	NC_005248	E. coli	γ proteobacteria	1.00	PL	NC_012124	S. enterica	γ proteobacteria	1.00
PL	NC_018666	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_010422	$S.\ enterica$	γ proteobacteria	1.00
PL	NC_018662	E. coli	$\gamma {\rm proteobacteria}$	1.00	PL	NC_002638	S. enterica	γ proteobacteria	1.00
PL	NC_013369	E. coli	γ proteobacteria	1.00	PL	NC_015965	S. enterica	γ proteobacteria	1.00
PL DI	NC_002142	E. coli E. coli	γ proteobacteria	1.00	PL	NC_005014	S. enterica	γ proteobacteria	1.00
гь PL	NC_009133 NC_017643	E. coli	vproteobacteria	1.00	PL PL	NC_019106	S. enterica S. enterica	vproteobacteria	1.00
PL	NC_017642	E. coli	γproteobacteria	1.00	PL	NC_011077	S. enterica	γproteobacteria	1.00
PL	NC_017640	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_011078	$S.\ enterica$	γ proteobacteria	1.00

PL	NC 019105	S. enterica	γ proteobacteria	1.00	PL	NC 015908	В.	bissettii	Spirochaetes	1.00
\mathbf{PL}	NC_010896	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_015909	В.	bissettii	Spirochaetes	1.00
PL	NC_019115	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_015910	В.	bissettii	Spirochaetes	1.00
\mathbf{PL}	NC_017720	S. enterica	γproteobacteria	1.00	\mathbf{PL}	NC_015911	В.	bissettii	Spirochaetes	1.00
\mathbf{PL}	NC_017054	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_015917	В.	bissettii	Spirochaetes	1.00
\mathbf{PL}	NC_019108	S. enterica	γ proteobacteria	1.00	PL	NC_015919	В.	bissettii	Spirochaetes	1.00
PL	NC_019109	S. enterica	γ proteobacteria	1.00	PL	NC_015922	В.	bissettii	Spirochaetes	1.00
PL	NC_006855	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_000948	В.	burgdorferi	Spirochaetes	1.00
PL	NC_019099	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_000949	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_019137	S. enterica	γ proteobacteria	1.00	PL	NC_000951	В.	burgdorferi	Spirochaetes	1.00
PL	NC_019002	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_000952	В.	burgdorferi	Spirochaetes	1.00
PL	NC_019001	S. enterica	γ proteobacteria	1.00	PL	NC_000953	В.	burgdorferi	Spirochaetes	1.00
PL	NC_017675	S. enterica	γ proteobacteria	1.00	PL	$NC_{-}000954$	В.	burgdorferi	Spirochaetes	1.00
PL	$NC_{-}016855$	S. enterica	γ proteobacteria	1.00	PL	NC_001850	В.	burgdorferi	Spirochaetes	1.00
PL	NC_019120	S. enterica	γ proteobacteria	1.00	PL	NC_001854	В.	burgdorferi	Spirochaetes	1.00
PL	NC_019125	S. enterica	γ proteobacteria	1.00	PL	NC_001857	Β.	burgdorferi	Spirochaetes	1.00
PL	NC_019126	S. enterica	γ proteobacteria	0.57	PL	NC_001903	Β.	burgdorferi	Spirochaetes	1.00
PL	NC_019122	S. enterica	γ proteobacteria	1.00	PL	NC_004971	В.	burgdorferi	Spirochaetes	1.00
PL	NC_019123	S. enterica	γ proteobacteria	1.00	PL	NC_011720	В.	burgdorferi	Spirochaetes	1.00
PL	NC_019131	S. enterica	γ proteobacteria	1.00	PL	NC_011722	В.	burgdorferi	Spirochaetes	1.00
PL	NC_010119	S. enterica	γproteobacteria	1.00	PL	NC_011724	В.	burgdorferi	Spirochaetes	1.00
PL	NC_016858	S. enterica	γproteobacteria	1.00	PL	NC_011731	В.	burgdorferi	Spirochaetes	1.00
PL	NC_013437	S. enterica	γproteobacteria	1.00	PL	NC_011735	В.	burgaorferi	Spirochaetes	1.00
PL	NC_010111	S. enterica	γproteobacteria	nan 1.00	PL	NC 011784	В. р	burgaorferi	Spirochaetes	1.00
FL DI	NC 010112	S. enterica	vproteobacteria	1.00		NC 011841	D. D	burgdorferi	Spirochaetes	1.00
PL	NC 011076	S. enterica	oproteobacteria	1.00	PL	NC 011843	D. B	burgdorferi	Spirochaetes	1.00
DI.	NC 011204	S. enterica	oproteobacteria	1.00	PL	NC 011845	D. B	burgdorferi	Spirochaetes	1.00
PL	NC 016864	S. enterica	vproteobacteria	1.00	PL	NC 011846	B.	burgdorferi	Spirochaetes	1.00
PL	NZ_CM00106:	S. enterica	γproteobacteria	1.00	PL	NC_011847	B.	buradorferi	Spirochaetes	1.00
PL	NZ_CM001154	1 S. enterica	γproteobacteria	1.00	PL	NC_011850	B.	burgdorferi	Spirochaetes	1.00
PL	NZ_CM00115	2 S. enterica	γ proteobacteria	1.00	PL	NC_011851	В.	buradorferi	Spirochaetes	1.00
PL	NC_002523	S. entomophila	γ proteobacteria	1.00	PL	NC_011853	В.	buradorferi	Spirochaetes	1.00
\mathbf{PL}	NC_002698	S. flexneri	γ proteobacteria	1.00	\mathbf{PL}	NC_011862	В.	burqdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_017319	S. flexneri	γproteobacteria	1.00	\mathbf{PL}	NC_011863	В.	burgdorferi	Spirochaetes	nan
\mathbf{PL}	NC_004851	S. flexneri	γ proteobacteria	1.00	\mathbf{PL}	NC_011866	В.	burgdorferi	Spirochaetes	1.00
PL	NC_013727	S. sonnei	γ proteobacteria	1.00	PL	NC_011876	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_016833	S. sonnei	γ proteobacteria	1.00	\mathbf{PL}	NC_011966	В.	burgdorferi	Spirochaetes	1.00
PL	NC_007385	S. sonnei	γ proteobacteria	1.00	\mathbf{PL}	NC_011967	В.	burgdorferi	Spirochaetes	1.00
PL	NC_010733	V. nigripulchritudo	γ proteobacteria	1.00	PL	NC_011968	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_015156	V. nigripulchritudo	γ proteobacteria	0.38	$_{\rm PL}$	NC_011970	В.	burgdorferi	Spirochaetes	1.00
PL	NC_010734	V. shilonii	γ proteobacteria	1.00	PL	NC_011974	Β.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_010114	Vibrio sp.	γ proteobacteria	0.56	PL	NC_011975	В.	burgdorferi	Spirochaetes	1.00
PL	NC_002120	Y. enterocolitica	γ proteobacteria	1.00	PL	NC_012104	В.	burgdorferi	Spirochaetes	1.00
PL	NC_004564	Y. enterocolitica	γproteobacteria	1.00	PL	NC_012106	В.	burgdorferi	Spirochaetes	1.00
PL	NC_005017	Y. enterocolitica	γproteobacteria	1.00	PL	NC_012107	В. Р	burgdorferi	Spirochaetes	1.00
PL	NC_010475	Y. enterocolitica	γproteobacteria	1.00	PL	NC_012110	В. р	burgaorferi	Spirochaetes	1.00
	NC 017565	Y enterocolitica	vproteobacteria	1.00		NC 012111	D. D	burgdorferi	Spirochaetes	1.00
PL	NC 008701	V enterocolitica	oproteobacteria	1.00	PL	NC 012112	D. B	burgdorferi	Spirochaetes	1.00
PL	NC 008122	Y nestis	vproteobacteria	1.00	PL	NC 012114	B.	burgdorferi	Spirochaetes	1.00
PL	NC 017153	Y nestis	vproteobacteria	1.00	PL	NC 012148	B.	burgdorferi	Spirochaetes	1.00
PL	NC_015056	Y. pestis	γproteobacteria	1.00	PL	NC_012149	B.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_017263	Y. pestis	γ proteobacteria	1.00	\mathbf{PL}	NC_012150	В.	burqdorferi	Spirochaetes	1.00
PL	NC_014017	Y. pestis	γ proteobacteria	1.00	\mathbf{PL}	NC_012151	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_010157	Y. pestis	γ proteobacteria	1.00	\mathbf{PL}	NC_012152	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_004836	Y. pestis	γ proteobacteria	1.00	\mathbf{PL}	NC_012153	В.	burgdorferi	Spirochaetes	1.00
PL	NC_004839	Y. pestis	γ proteobacteria	1.00	\mathbf{PL}	NC_012154	В.	burgdorferi	Spirochaetes	1.00
PL	NC_003131	Y. pestis	γ proteobacteria	1.00	PL	NC_012157	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_017157	Y. pestis	γ proteobacteria	1.00	PL	NC_012158	В.	burgdorferi	Spirochaetes	1.00
PL	NC_005813	Y. pestis	γ proteobacteria	1.00	PL	NC_012159	В.	burgdorferi	Spirochaetes	1.00
PL	NC_009595	Y. pestis	γ proteobacteria	1.00	PL	NC_012160	В.	burgdorferi	Spirochaetes	1.00
PL	NC_009377	Y. pestis	γ proteobacteria	1.00	PL	NC_012161	В.	burgdorferi	Spirochaetes	1.00
PL	NC_010635	Y. pseudotuberculosis	γproteobacteria	1.00	PL	NC_012165	В.	burgdorferi	Spirochaetes	1.00
PL DI	NC 006152	1. pseudotuberculosis	γproteopacteria	1.00		NC 012172	В. Р	ourgaorferi	Spirochaetes	1.00
ГL	NC_000133	1. pseudotaberculosis	yproteobacteria	1.00	PL	NC 012174	D. B	burgdorferi	Spirochaetes	1.00
clust	er : 8 stability	measure : 0.985116			PL	NC 012178	B.	burgdorferi	Spirochaetes	1.00
PL.	NC 008273	B afzelij	Spirochaetes	1.00	PL	NC 012187	B.	burgdorferi	Spirochaetes	1.00
PL	NC_008274	B. afzelii	Spirochaetes	1.00	PL	NC_012193	В.	burgdorferi	Spirochaetes	1.00
PL	NC_008564	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_012194	В.	burqdorferi	Spirochaetes	1.00
PL	NC_011647	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_012199	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_011648	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_012229	В.	burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_011650	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_012235	В.	burgdorferi	Spirochaetes	nan
PL	NC_011651	B. afzelii	Spirochaetes	1.00	PL	NC_012238	В.	burgdorferi	Spirochaetes	1.00
PL	NC_011786	B. afzelii	Spirochaetes	1.00	PL	NC_012240	В.	burgdorferi	Spirochaetes	0.50
\mathbf{PL}	NC_011790	B. afzelii	Spirochaetes	1.00	PL	NC_012241	В.	burgdorferi	Spirochaetes	1.00
PL	NC_011791	B. afzelii	Spirochaetes	1.00	PL	NC_012244	В.	burgdorferi	Spirochaetes	1.00
PL	NC_017224	B. afzelii	Spirochaetes	1.00	PL	NC_012248	В.	burgdorferi	Spirochaetes	1.00
PL	NC_017225	B. afzelii	Spirochaetes	1.00	PL	NC_012250	B.	burgdorferi	Spirochaetes	1.00
PL	NC_017226	B. afzelii	Spirochaetes	1.00	PL	NC_012251	В.	ourgdorferi	Spirochaetes	1.00
PL	NC 017227	D. ajzein P. afzelii	Spirochaetes	1.00		NC 012253	В. Р	ourgaorferi	Spirochaetes	1.00
PL DI	NC 017228	D. ajzenn B. afzelić	Spirochaetes	1.00	PL DI	NC 012254	В. Р	burgdorferi	Spirochaetes	1.00
PL DI	NC_{017229}	D. ajzeni B. afzelij	Spirochaetes	1.00	PL DI	NC 012255	<i>В</i> . В	burgdorferi	Spirochaetes	1.00
гь рт	NC 017221	D. ujzeni B. afzelić	Spirochaetes	1.00		NC 012258	ы. R	huradorferi	Spirochaetes	1.00
E 14	110-01/231	D. ujzenn D. afzolii	Spirochaetes	1.00	PI.	NC_012259	B. R	burgdorferi	Spirochaetes	1.00
рт	NC 017934	1) 1112F112		1 1.00		NG 010001	D.		Sphochaetes	1.00
PL PL	NC_017234 NC_017236	B. afzelii	Spirochaetes	1.00	PL 1	NC_012261	B	burgdorferi	Spirochaetes	1.00
PL PL PL	NC_017234 NC_017236 NC_017241	B. afzelii B. afzelii B. afzelii	Spirochaetes	1.00	PL PL	NC_012261 NC_012262	B. B	burgdorferi burgdorferi	Spirochaetes	1.00
PL PL PL PL PI.	NC_017234 NC_017236 NC_017241 NC_015903	B. afzelii B. afzelii B. bissettii	Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00	PL PL PL	NC_012261 NC_012262 NC_012263	В. В. В.	burgdorferi burgdorferi burgdorferi	Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00
PL PL PL PL PL PL	NC_017234 NC_017236 NC_017241 NC_015903 NC_015905	B. afzelii B. afzelii B. bissettii B. bissettii	Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 1.00	PL PL PL PL	NC_012261 NC_012262 NC_012263 NC_012266	В. В. В. В.	burgdorferi burgdorferi burgdorferi burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL	NC_017234 NC_017236 NC_017241 NC_015903 NC_015905 NC_015906	B. ajzetii B. afzelii B. bissettii B. bissettii B. bissettii	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL	NC_012261 NC_012262 NC_012263 NC_012266 NC_012499	В. В. В. В.	burgdorferi burgdorferi burgdorferi burgdorferi burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 1.00 0.64
PL PL PL PL PL PL PL	NC_017234 NC_017236 NC_017241 NC_015903 NC_015905 NC_015906 NC_015907	D. afzelii B. afzelii B. ofissettii B. bissettii B. bissettii B. bissettii	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$ \begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ \end{array} $	PL PL PL PL PL PL	NC_012261 NC_012262 NC_012263 NC_012266 NC_012499 NC_012501	В. В. В. В. В.	burgdorferi burgdorferi burgdorferi burgdorferi burgdorferi burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$ \begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \\ 0.64 \\ 1.00 \end{array} $

DT	NG 010502		g : l	0.64	DI	NC 017705	D · ··	a :	1 1 00
PL DI	NC_012505	B. burgaorjeri	Spirochaetes	0.64	PL	NC_017725	Б. garinii	Spirochaetes	1.00
PL	NC_012504	B. burgdorferi	Spirochaetes	1.00	PL	NC_017804	$B. \ garinii$	Spirochaetes	0.64
PL	$NC_{-}012505$	B. burgdorferi	Spirochaetes	1.00	PL	NC_011246	B. recurrentis	Spirochaetes	1.00
\mathbf{PL}	NC_012509	B. burgdorferi	Spirochaetes	1.00	PL	NC_011252	B. recurrentis	Spirochaetes	1.00
PL	NC_012510	B. burgdorferi	Spirochaetes	1.00	PL	NC_011253	B. recurrentis	Spirochaetes	1.00
PL.	NC 012512	B huradorferi	Spirochaetes	1.00	PL.	NC 011255	B recurrentis	Spirochaetes	1.00
DI	NC 012512	B. burgdorfert	Sphoenaetes	1.00		NC 011260	B. recurrentia	Spirochaetes	1.00
PL DI	NC_012515	B. burgaonjeni	Spirochaetes	1.00	PL DI	NC_011200	B. recurrentis	Spirochaetes	1.00
PL	NC_012514	B. burgdorferi	Spirochaetes	1.00	PL	NC_012132	B. spielmanii	Spirochaetes	1.00
PL	NC_012515	B. burgdorferi	Spirochaetes	1.00	PL	NC_012164	B. spielmanii	Spirochaetes	1.00
\mathbf{PL}	NC_012516	B. burgdorferi	Spirochaetes	1.00	PL	NC_012173	B. spielmanii	Spirochaetes	1.00
PL.	NC 012517	B huradorferi	Spirochaetes	1.00	PL.	NC 012181	B spielmanii	Spirochaetes	1.00
DI	NC 012519	B. burgaorferi	Spirochaetes	1.00		NC 012101	B. spicinianti	Spirochaetes	1.00
PL	NC_012518	B. burgaonjen	Spirochaetes	1.00	PL	NC_012128	B. valaisiana	Spirochaetes	1.00
PL	NC_013128	B. burgdorferi	Spirochaetes	1.00	PL	NC_012129	B. valaisiana	Spirochaetes	1.00
\mathbf{PL}	NC_013129	B. burgdorferi	Spirochaetes	1.00	PL	NC_012130	B. valaisiana	Spirochaetes	1.00
PL	NC_013130	B. burgdorferi	Spirochaetes	1.00	PL	NC_012131	B. valaisiana	Spirochaetes	1.00
\mathbf{PL}	NC 017393	B. buradorferi	Spirochaetes	1.00	PL	NC 012133	B. valaisiana	Spirochaetes	1.00
DI	NC 017204	P humadomfomi	Spirachaotac	1.00	 DI	NC 012166	P. walaisiana	Spirochaster	1.00
	NC_017394	B. burguorjeri	Spirochaetes	1.00	FL	NC_012100	D. valaisiana	Sphochaetes	1.00
PL	NC_017395	B. burgdorferi	Spirochaetes	1.00	PL	NC_012169	B. valaisiana	Spirochaetes	1.00
PL	NC_017396	B. burgdorferi	Spirochaetes	1.00	PL	NC_012177	B. valaisiana	Spirochaetes	1.00
\mathbf{PL}	NC_017397	B. burgdorferi	Spirochaetes	1.00	PL	NC_012185	B. valaisiana	Spirochaetes	1.00
\mathbf{PL}	NC_017398	B. buradorferi	Spirochaetes	1.00	PL	NC_012230	Borrelia sp.	Spirochaetes	1.00
PL.	NC 017401	B huradorferi	Spirochaetes	1.00	PI.	NC 012234	Borrelia SD	Spirochaetes	1.00
DI	NC 017402	B. burgaorferi	Spirochaetes	1.00		NC 019959	Bornelia an	Spirochaetes	1.00
PL	NC_017402	B. burgaonjen	Spirochaetes	1.00	PL	NC_012252	Borrelia sp.	spirochaetes	1.00
PL	NC_017406	B. burgdorferi	Spirochaetes	1.00	PL	NC_012265	Borrelia sp.	Spirochaetes	1.00
PL	NC_017407	B. burgdorferi	Spirochaetes	1.00	PL	NC_012267	Borrelia sp.	Spirochaetes	1.00
\mathbf{PL}	NC_017409	B. burgdorferi	Spirochaetes	1.00	PL	NC_010844	L. biflexa	Spirochaetes	1.00
PL.	NC 017415	B huradorferi	Spirochaetes	1.00	PL	NC 010846	L biflera	Spirochaetes	1.00
DI	NC 017416	D. burgaorfort	Spirochaetee	0.72	12	1101010010	El objectad	Sphoendores	1.00
PL DI	NC_017410	B. burgaonjeni	Spirochaetes	0.73					
PL	NC_017420	B. burgdorferi	Spirochaetes	1.00	cluster	: 9 stability n	ieasure : 0.973547		
PL	NC_017422	B. burgdorferi	Spirochaetes	1.00	CHR	NC_009850	A. butzleri	ϵ proteobacteria	1.00
\mathbf{PL}	NC_017423	B. burgdorferi	Spirochaetes	1.00	CHR	NC_017187	A. butzleri	ϵ proteobacteria	1.00
\mathbf{PL}	NC 017424	B. buradorferi	Spirochaetes	1.00	СНВ	NC 014166	A nitrofiailis	eproteobacteria	1.00
DI	NC 017425	P humadomfori	Spirochaotos	1.00	CHR	NC 017100	A. nutrojiguis	epioteobacteria	1.00
PL DI	NC_017425	B. burgaonjen	Spirochaetes	1.00	CHR	NC_017192	Arcobacter sp.	eproteobacteria	1.00
PL	NC_017426	B. burgdorferi	Spirochaetes	1.00	PL	NC_009795	C. concisus	€proteobacteria	1.00
\mathbf{PL}	NC_017427	B. burgdorferi	Spirochaetes	1.00	CHR	NC_009802	C. concisus	ϵ proteobacteria	1.00
\mathbf{PL}	NC_017428	B. burgdorferi	Spirochaetes	1.00	CHR	NC_009715	C. curvus	εproteobacteria	1.00
PL.	NC 018979	B huradorferi	Spirochaetes	1.00	СНВ	NC 008500	C fatus	sproteobacteria	1.00
DI	NC 018080	B. burgaorferi	Spirochaetes	1.00	CIII	NC 000714	C. Jetus	eproteobacteria	1.00
PL	NC_018980	B. burgaonjen	Spirochaetes	1.00	CHR	NC_009714	C. nominis	eproteobacteria	1.00
PL	NC_018981	B. burgdorferi	Spirochaetes	1.00	CHR	NC_002163	C. jejuni	ϵ proteobacteria	1.00
\mathbf{PL}	NC_018982	B. burgdorferi	Spirochaetes	1.00	CHR	NC_003912	C. jejuni	ϵ proteobacteria	1.00
PL	NC_018983	B. buradorferi	Spirochaetes	1.00	PL	$NC_{-}005012$	C. jejuni	eproteobacteria	1.00
PL.	NC 018989	B huradorferi	Spirochaetes	1.00	СНВ	NC 008787	C. jejunj	eproteobacteria	1.00
DI	NC 018002	B. burgabijert	Sphoenaetes	0.64	CIII	NC 000707		eproteobacteria	1.00
PL	NC-018993	B. burgaorferi	Spirochaetes	0.64	CHR	NC_009707	C. jejuni	eproteobacteria	1.00
PL	NC_019003	B. burgdorferi	Spirochaetes	1.00	CHR	NC_009839	C. jejuni	ϵ proteobacteria	1.00
PL	NC_019004	B. burgdorferi	Spirochaetes	1.00	CHR	NC_014802	C. jejuni	ϵ proteobacteria	1.00
PL	NC_019005	B. burgdorferi	Spirochaetes	1.00	CHR	NC_017279	C. jejuni	ϵ proteobacteria	1.00
\mathbf{PL}	NC 019006	B. buradorferi	Spirochaetes	1.00	CHB	NC 017280	C jejuni	eproteobacteria	1.00
DI	NC 017779	B. ourguorjert	Sphoenaetes	1.00	CIII	NC 017200		eproteobacteria	1.00
	NC_017778	B. crociaurae	Spirochaetes	1.00	CHR	NC_017281	C. jejuni	εproteobacteria	1.00
PL	NC_017780	B. crocidurae	Spirochaetes	1.00	CHR	NC_018521	C. jejuni	ϵ proteobacteria	1.00
PL	NC_017782	B. crocidurae	Spirochaetes	1.00	CHR	NC_018709	C. jejuni	ϵ proteobacteria	nan
\mathbf{PL}	NC_017783	B. crocidurae	Spirochaetes	1.00	CHR	NC_012039	C. lari	εproteobacteria	1.00
PL.	NC 017786	B crocidurae	Spirochaetes	1.00	СНВ	NC 008220	H acinonuchis	sproteobacteria	1.00
	NC 017797	D. crociaurae	Spirochaetes	1.00		NC 015670	II. acinonychis	eproteobacteria	1.00
PL	NC_017787	B. crociaurae	Spirochaetes	1.00	PL	NC_015670	H. bizzozeronii	eproteobacteria	0.95
PL	NC_017788	B. crocidurae	Spirochaetes	1.00	CHR	NC_015674	H. bizzozeronii	ϵ proteobacteria	1.00
\mathbf{PL}	NC_017789	B. crocidurae	Spirochaetes	1.00	CHR	NC_017735	H. cetorum	ϵ proteobacteria	1.00
PL	NC_017794	B. crocidurae	Spirochaetes	1.00	CHB	NC 017737	H cetorum	eproteobacteria	1.00
PL.	NC 017796	B crocidurae	Spirochaetes	1.00	CHB	NC 017761	H cinaedi	sproteobactoria	1.00
	NC 017707	D. crociaarac	Gindentees	1.00	CIII	NG 01/010	II. cinucui	eproteobacteria	1.00
PL	NC_017797	B. crocidurae	Spirochaetes	1.00	CHR	NC_014810	H. felis	ϵ proteobacteria	1.00
PL	NC_017801	B. crocidurae	Spirochaetes	1.00	CHR	NC_004917	H. hepaticus	ϵ proteobacteria	1.00
\mathbf{PL}	NC_{017802}	B. crocidurae	Spirochaetes	1.00	CHR	NC_013949	H. mustelae	ϵ proteobacteria	1.00
PL	NC_017811	B. crocidurae	Spirochaetes	1.00	CHR	NC_000915	H. pylori	ϵ proteobacteria	1.00
PL	NC_017812	B. crocidurae	Spirochaetes	1.00	CHB	NC.000921	H. pulori	eproteobacteria	1.00
рт	NC 01791F	B crocidurac	Spirochaster	1.00	CUD	NC 0000021	U mulori	cproteobacteria	1.00
L D	NCL017815	D. Crociaurae	Spirochaetes	1.00	CHR	NG 010005	n. pylori	ϵ proteobacteria	1.00
PL	NG_017817	В. crocidurae	Spirochaetes	1.00	CHR	NC_010698	H. pylori	ϵ proteobacteria	1.00
PL	NC_017818	B. crocidurae	Spirochaetes	1.00	CHR	NC_011333	H. pylori	ϵ proteobacteria	1.00
\mathbf{PL}	NC_017819	B. crocidurae	Spirochaetes	1.00	CHR	NC_011498	H. pylori	ϵ proteobacteria	1.00
PL	NC_017821	B. crocidurae	Spirochaetes	1.00	CHR	NC_012973	H. pylori	ϵ proteobacteria	1.00
PL	NC_011226	B. duttonii	Spirochaetes	1.00	CHB	NC 014256	H. pulori	eproteobacteria	1.00
рт	NC 011245	B duttonii	Spirochasto	1 00	CUD	NC 014FFF	H malomi	cprotoch-start	1.00
	NC 011240	D. autionni D. J., C	Sprochaetes	1.00	CHR	NC 014555	11. pytori	eproteobacteria	1.00
rL T	NG_011247	 auttonii 	Spirochaetes	1.00	CHR	INC_014560	н. pylori	ϵ proteobacteria	1.00
PL	NC_011248	B. duttonii	Spirochaetes	1.00	CHR	NC_017063	H. pylori	ϵ proteobacteria	1.00
PL	NC_011249	B. duttonii	Spirochaetes	1.00	CHR	NC_017354	H. pylori	ϵ proteobacteria	1.00
\mathbf{PL}	NC_011250	B. duttonii	Spirochaetes	nan	CHB	NC 017355	H nulori	eproteobacteria	1.00
рт	NC 011251	B duttonii	Spirochaetec	1.00	СПР	NC 017257	H mulori	eprotochactoria	1.00
л. рт	NC 011254	B duttonii	Chinochaetes	1.00		NC 017959	U mul	epioteobacteria	1.00
	NG 011250	D. autionni	Sprochaetes	1.00	Снк	NG_017398	11. pytori	eproteobacteria	1.00
PL	NC_011257	B. duttonii	Spirochaetes	1.00	CHR	NC_017359	H. pylori	ϵ proteobacteria	1.00
PL	NC_011259	B. duttonii	Spirochaetes	1.00	CHR	NC_017360	H. pylori	ϵ proteobacteria	1.00
PL	NC_011261	B. duttonii	Spirochaetes	1.00	CHR	NC_017361	H. pylori	ϵ proteobacteria	1.00
PL	NC_011262	B. duttonii	Spirochaetes	1.00	CHB	NC.017362	H. pulori	eproteobacteria	1.00
рт	NC 011264	B duttonii	Spirochasto	1.00	CUD	NC 017965	H palomi	cprotoch-start	1.00
	NC 011204	D. duitofill	Spirochaetes	1.00	CHR	NC-017305	11. pytori	eproteobacteria	1.00
PL	NC_011265	B. duttonii	Spirochaetes	1.00	CHR	NC_017366	H. pylori	ϵ proteobacteria	1.00
PL	NC_006128	B. garinii	Spirochaetes	1.00	CHR	NC_017367	H. pylori	ϵ proteobacteria	1.00
PL	NC_006129	B. garinii	Spirochaetes	1.00	CHR	NC_017368	H. pylori	ϵ proteobacteria	1.00
рт	NC 011840	B aarinii	Spirochaetec	1.00	СПВ	NC 017271	H mulori	eprotochactoria	1.00
л. рт	NC 011040	B garinii	Chinochaetes	1.00		NC 017979	U mul	epioteobacteria	1.00
	NC 011042	D. guriniti	Spirochaetes	1.00	CHR	NG_01/3/2	11. pytori	eproteobacteria	1.00
PL	NC_011844	в. garınıı	Spirochaetes	1.00	CHR	NC_017374	H. pylori	ϵ proteobacteria	1.00
PL	NC_011848	B. garinii	Spirochaetes	1.00	CHR	NC_{017375}	H. pylori	ϵ proteobacteria	1.00
\mathbf{PL}	NC_011856	B. garinii	Spirochaetes	1.00	CHR	NC_017376	H. pylori	ϵ proteobacteria	1.00
PL.	NC.011859	B. garinii	Spirochaetes	1.00	CHR	NC 017378	H. nulori	eproteobacteria	1.00
PT	NC 011860	B garinii	Spirochaetes	1 00	СПР	NC 017970	H mulori	cprotochacteria	1.00
DT	NC 011000	D. garinii D. garinii	C_:l	1.00	CIR	NC 017001	II. pytori	eproteobacteria	1.00
PL	NG_011869	в. garınıı	Spirochaetes	nan	CHR	NC_017381	H. pylori	ϵ proteobacteria	1.00
	NC_011875	B. garinii	Spirochaetes	1.00	CHR	NC_017382	H. pylori	ϵ proteobacteria	1.00
PL	10.00	B				MCL OIE000			

	NG OFFER	TT T ·		1 1 00	DI	NG 010001		D	1 00
CHR	NC_017733	H. pylori	eproteobacteria	1.00	PL	NC_012001	M. caseolyticus	Bacilli	1.00
CHR	NC_017739	H. pylori	eproteobacteria	1.00	PL	NC_012002	M. caseolyticus	Bacilli	1.00
CHR	NC_017740	H. pylori	eproteobacteria	1.00	PL	NZ_CM00139	9 O. kitanarae	Bacilli	1.00
CHR	NC_017741	H. pylori	€proteobacteria	1.00	PL	NC_002517	S. aureus	Bacilli	1.00
CHR	NC_017742	H. pylori	ϵ proteobacteria	1.00	PL	NC_002774	S. aureus	Bacilli	0.77
PL	NC_017919	H. pylori	ϵ proteobacteria	0.68	PL	NC_003140	S. aureus	Bacilli	1.00
CHR	NC_017926	H. pylori	ϵ proteobacteria	1.00	PL	NC_003265	S. aureus	Bacilli	1.00
CHR	NC_018937	H. pylori	ϵ proteobacteria	1.00	PL	NC_005011	S. aureus	Bacilli	1.00
CHR	NC_018938	H. pylori	ϵ proteobacteria	1.00	PL	NC_005024	S. aureus	Bacilli	0.06
CHR	NC_018939	H. pylori	ϵ proteobacteria	1.00	PL	$NC_{-}005054$	S. aureus	Bacilli	0.50
CHR	NC_012115	N. profundicola	ϵ proteobacteria	1.00	PL	NC_005127	S. aureus	Bacilli	1.00
CHR	NC_014935	N. salsuginis	ϵ proteobacteria	1.00	PL	NC_005951	S. aureus	Bacilli	1.00
CHR	NC_009662	Nitratiruptor sp.	ϵ proteobacteria	1.00	PL	NC_007792	S. aureus	Bacilli	0.56
CHR	NC_014506	S. autotrophica	eproteobacteria	1.00	PL	NC_007931	S. aureus	Bacilli	1.00
CHR	NC_018002	S. barnesii	eproteobacteria	1.00	PL	NC_009477	S. aureus	Bacilli	1.00
CHR	NC 013512	S. delevianum	eproteobacteria	1.00	PL	NC 009619	S. aureus	Bacilli	1.00
CHB	NC 007575	S denitrificans	eproteobacteria	1.00	PL.	NC 010063	S aureus	Bacilli	1.00
CHB	NC 014762	S kujiense	eproteobacteria	1.00	PL.	NC 010066	S. aureus	Bacilli	1 00
CHR	NC 009663	Sulfarovorum SD	eproteobacteria	1.00	PL	NC 010077	S aureus	Bacilli	1.00
CHR	NC 005000	W succinogenes	eproteobacteria	1.00	PL	NC 010279	S. aureus	Bacilli	0.38
DI	NC 012280	Restorium co	Human gut	0.12	P L DI	NC 010410	S. aureus	Bacilli	1.00
ГL	10101010200	Bucterium sp.	iiuman gut	0.12		NC 011522	S. uureus	Daciiii	1.00
	10 -t - L:1:t					NC 012547	S. aureus	Daciiii Daciiii	1.00
cluster :	NG accord	leasure : 0.980952			PL	NC_0120247	S. aureus	Bacilli	1.00
CHR	NC_009925	A. marina	Chroobacteria	nan	PL	NC_013034	S. aureus	Bacilli	1.00
CHR	NC_013771	Cyanothece UCYN-A	Chroobacteria	1.00	PL	NC_013289	S. aureus	Bacilli	1.00
CHR	NC_010546	Cyanothece sp.	Chroobacteria	1.00	PL	NC_013290	S. aureus	Bacilli	1.00
CHR	NC_011726	Cyanothece sp.	Chroobacteria	1.00	PL	NC_013292	S. aureus	Bacilli	1.00
CHR	NC_011729	Cyanothece sp.	Chroobacteria	1.00	PL	NC_013293	S. aureus	Bacilli	1.00
CHR	NC_011884	Cyanothece sp.	Chroobacteria	1.00	PL	NC_013294	S. aureus	Bacilli	1.00
CHR	NC_013161	Cyanothece sp.	Chroobacteria	1.00	PL	NC_013296	S. aureus	Bacilli	1.00
CHR	NC_014501	Cyanothece sp.	Chroobacteria	1.00	PL	NC_013298	S. aureus	Bacilli	1.00
CHR	NC_010296	M. aeruginosa	Chroobacteria	nan	PL	NC_013299	S. aureus	Bacilli	1.00
CHR	NC_006576	S. elongatus	Chroobacteria	1.00	PL	NC_013301	S. aureus	Bacilli	1.00
CHR	NC_007604	S. elongatus	Chroobacteria	1.00	PL	NC_013303	S. aureus	Bacilli	1.00
CHR	NC_005070	Synechococcus sp.	Chroobacteria	1.00	PL	NC_013304	S. aureus	Bacilli	1.00
CHR	NC_007513	Synechococcus sp.	Chroobacteria	1.00	PL	NC_013313	S. aureus	Bacilli	1.00
CHR	NC_007516	Synechococcus sp.	Chroobacteria	1.00	PL	NC_013318	S. aureus	Bacilli	1.00
CHR	NC_007775	Synechococcus sp.	Chroobacteria	1.00	PL	NC_013319	S. aureus	Bacilli	1.00
CHR	NC_007776	Sunechococcus sp.	Chroobacteria	1.00	PL	NC_013320	S. aureus	Bacilli	0.68
CHR	NC 008319	Sunechococcus sp.	Chroobacteria	1.00	PL	NC_013321	S. aureus	Bacilli	1.00
CHB	NC 009481	Sunechococcus sp	Chroobacteria	1.00	PL	NC_013322	S. aureus	Bacilli	1.00
CHR	NC 009482	Sunechococcus sp.	Chroobacteria	1.00	PL	NC_013323	S. aureus	Bacilli	1.00
CHB	NC 010475	Supechococcus sp	Chroobacteria	1.00	PL	NC 013324	S aureus	Bacilli	1.00
CHB	NC 000911	Supechocustis sp.	Chroobacteria	1.00	PL	NC 013326	S. aureus	Bacilli	1 00
CHR	NC 017038	Synechocystic sp.	Chroobacteria	1.00	PL	NC 013330	S. aureus	Bacilli	1 00
CHR	NC 017038	Synechocystis sp.	Chroobacteria	1.00	PL	NC 013331	S. aureus	Bacilli	1.00
CHR	NC 017052	Synechocystis sp.	Chroobacteria	1.00	DI DI	NC 013332	S aureus	Bacilli	1.00
CHR	NC 017052	Synechocystis sp.	Chroobacteria	1.00	DI	NC 012222	S. aureus	Pacilli	1.00
CHR	NC_01/2//	Synechocystis sp.	Chroobacteria	1.00		NC 012224	S. aureus	Daciiii Daciiii	1.00
CHR	NC_004113	1. elongatus	Chroobacteria	1.00	PL	NC_015554	S. aureus	Bacilli	1.00
CHR	NC_008312	T. erythraeum	Chroobacteria	1.00	PL	NC_013335	S. aureus	Bacilli	1.00
CHR	NC_005125	G. violaceus	Gloeobacteria	1.00	PL	NC_013337	S. aureus	Bacilli	1.00
CHR	NC_007413	A. variabilis	Homogonae	1.00	PL	NC_013338	S. aureus	Bacilli	0.50
CHR	NC_019427	Anabaena sp.	Homogonae	1.00	PL	NC_013339	S. aureus	Bacilli	0.32
RECE	NC_019439	Anabaena sp.	Homogonae	1.00	PL	NC_013340	S. aureus	Bacilli	1.00
CHR	NC_014248	N. azollae	Homogonae	1.00	PL	NC_013342	S. aureus	Bacilli	0.50
CHR	NC_010628	N. punctiforme	Homogonae	1.00	PL	NC_013343	S. aureus	Bacilli	0.56
CHR	NC_014248	N. azollae	Homogonae	1.00	PL	NC_013344	S. aureus	Bacilli	1.00
CHR	NC_003272	Nostoc sp.	Homogonae	1.00	PL	NC_013347	S. aureus	Bacilli	nan
CHR	NC_005042	P. marinus	Prochlorales	1.00	PL	NC_013348	S. aureus	Bacilli	1.00
CHR	NC_005071	P. marinus	Prochlorales	1.00	PL	NC_013349	S. aureus	Bacilli	1.00
CHR	NC_005072	P. marinus	Prochlorales	1.00	PL	NC_013351	S. aureus	Bacilli	1.00
CHR	NC_007335	P. marinus	Prochlorales	1.00	PL	NC_013352	S. aureus	Bacilli	1.00
CHR	NC_007577	P. marinus	Prochlorales	1.00	PL	NC_013371	S. aureus	Bacilli	1.00
CHR	NC_008816	P. marinus	Prochlorales	1.00	PL	NC_013372	S. aureus	Bacilli	1.00
CHR	NC_008817	P. marinus	Prochlorales	1.00	PL	NC_013374	S. aureus	Bacilli	1.00
CHR	NC_008819	D '	Prochlorales	1 1 00	PL		a	Bacilli	0.64
CHB		P. marinus	ribemorales	1.00		NC_013377	S. aureus	1	
Onn	NC_008820	P. marinus P. marinus	Prochlorales	1.00	PL	NC_013377 NC_013378	S. aureus S. aureus	Bacilli	0.43
CHR	NC_008820 NC_009091	P. marinus P. marinus P. marinus	Prochlorales Prochlorales	1.00 1.00 1.00	PL PL	NC_013377 NC_013378 NC_013379	5. aureus 5. aureus 5. aureus	Bacilli Bacilli	$0.43 \\ 1.00$
CHR CHR	NC_008820 NC_009091 NC_009840	P. marinus P. marinus P. marinus P. marinus	Prochlorales Prochlorales Prochlorales	1.00 1.00 1.00	PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380	S. aureus S. aureus S. aureus S. aureus	Bacilli Bacilli Bacilli	$0.43 \\ 1.00 \\ 1.00$
CHR CHR CHR	NC_008820 NC_009091 NC_009840 NC_009976	P. marinus P. marinus P. marinus P. marinus P. marinus	Prochlorales Prochlorales Prochlorales Prochlorales	$ \begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array} $	PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381	S. aureus S. aureus S. aureus S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli	$0.43 \\ 1.00 \\ 1.00 \\ 1.00$
CHR CHR CHR	NC_008820 NC_009091 NC_009840 NC_009976	P. marinus P. marinus P. marinus P. marinus P. marinus	Prochlorales Prochlorales Prochlorales Prochlorales	$ \begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array} $	PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382	S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli	$0.43 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00$
CHR CHR CHR cHR	NC_008820 NC_009091 NC_009840 NC_009976 : 11 stability n	P. marinus P. marinus P. marinus P. marinus P. marinus neasure : 0.928523	Prochlorales Prochlorales Prochlorales Prochlorales	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00 \end{array} $	PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013383	S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$0.43 \\ 1.00 \\ $
CHR CHR CHR CHR cluster : PL	NC_008820 NC_009091 NC_009840 NC_009976 : 11 stability n NC_011332	P. marinus P. marinus P. marinus P. marinus P. marinus neasure : 0.928523 B. bifidum	Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013383 NC_013384	S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$0.43 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 $
CHR CHR CHR CHR <u>cluster</u> : PL PL	NC_008820 NC_009091 NC_009840 NC_009976 : 11 stability n NC_011332 NC_015053	P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum	Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013383 NC_013384 NC_013387	S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array}$
CHR CHR CHR CHR CHR PL PL PL PL	NC.008820 NC.009091 NC.009840 NC.009976 11 stability n NC.011332 NC.015053 NC.010857	P. marinus P. marinus P. marinus P. marinus P. marinus B. bifidum B. longum B. longum	Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013383 NC_013384 NC_013387 NC_013388	S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array}$
CHR CHR CHR CHR PL PL PL PL PL	NC_008820 NC_009091 NC_009840 NC_009976 : 11 stability m NC_011322 NC_010857 NC_017222	P. marinus P. marinus P. marinus P. marinus P. marinus B. bifidum B. longum B. longum B. longum	Actinobacteridae Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL	NC.013377 NC.013378 NC.013379 NC.013380 NC.013381 NC.013382 NC.013384 NC.013384 NC.013387 NC.013388 NC.013389	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\end{array}$
CHR CHR CHR CHR PL PL PL PL PL PL PL	NC_008820 NC_009091 NC_009840 NC_009976 : 11 stability m NC_011332 NC_015053 NC_010857 NC_017222 NC_004978	P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum B. longum B. longum	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013383 NC_013384 NC_013387 NC_013389 NC_013390	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\end{array}$
CHR CHR CHR PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009940 NC.009976 : 11 stability n NC.01322 NC.015053 NC.010857 NC.017222 NC.004978 NC.004770	P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum B. longum B. longum B. longum	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013384 NC_013384 NC_013387 NC_013388 NC_013389 NC_013390 NC_013393	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\end{array}$
CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL	NC_008820 NC_009091 NC_009840 NC_009976 : 11 stability n NC_011332 NC_01553 NC_010857 NC_017222 NC_004978 NC_004978 NC_004770 NC_010164	P. marinus P. marinus P. marinus P. marinus P. marinus B. bifidum B. bifidum B. longum B. longum B. longum B. longum B. longum Bifidobacterium sp.	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013383 NC_013384 NC_013387 NC_013388 NC_013389 NC_013390 NC_013393 NC_013550	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array}$
CHR CHR CHR CHR PL PL PL PL PL PL PL PL	NC_008820 NC_009091 NC_009976 : 11 stability n NC_01322 NC_010857 NC_017222 NC_004978 NC_004978 NC_004770 NC_010394	P. marinus P. marinus P. marinus P. marinus P. marinus B. bifidum B. longum B. longum B. longum B. longum B. longum B. longum M. abscessus	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.54	PL PL PL PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013384 NC_013384 NC_013384 NC_013389 NC_013390 NC_013390 NC_013350 NC_013653	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\end{array}$
CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 11 stability n NC.01332 NC.015053 NC.010857 NC.017222 NC.004978 NC.004770 NC.010164 NC.010394 NC 010604	P. marinus P. marinus P. marinus P. marinus P. marinus B. bifidum B. longum B. longum B. longum B. longum B. longum B. longum B. longum M. abscessus M. marinum	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.63\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013384 NC_013384 NC_013387 NC_013389 NC_013390 NC_013390 NC_013390 NC_013550 NC_014369	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\end{array}$
CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009940 NC.009976 : 11 stability n NC.015053 NC.015053 NC.015053 NC.010857 NC.017222 NC.004978 NC.004770 NC.010164 NC.010394 NC.013793	P. marinus P. marinus P. marinus P. marinus P. marinus B. bifidum B. longum B. longum	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.63\\ 0.79\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL PL PL PI.	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013384 NC_013384 NC_013387 NC_013388 NC_013389 NC_013390 NC_013550 NC_013653 NC_016942	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.43 1.00
CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_008820 NC_009091 NC_009976 : 11 stability n NC_011332 NC_01553 NC_010857 NC_017222 NC_004978 NC_004978 NC_004978 NC_010164 NC_010394 NC_01664 NC_013793 NC_00512	P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum B. longum B. longum B. longum Bifidobacterium sp. M. abscessus M. marinum B. pseudofirmus F. forecire	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.63\\ 0.79\\ 1.00\\ 1.00\\ 0.54\\ 0.63\\ 0.79\\ 1.00\\ 0.54\\ 0.63\\ 0.79\\ 0.54\\ 0.63\\ 0.79\\ 0.54\\ 0.63\\ 0.79\\ 0.54\\ 0.63\\ 0.79\\ 0.54\\ 0.63\\ 0.79\\ 0.54\\ 0.54\\ 0.54\\ 0.56\\ 0.79\\ 0.56\\ 0.56\\ 0.79\\ 0.56\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013383 NC_013384 NC_013384 NC_013389 NC_013389 NC_013390 NC_013390 NC_013653 NC_013653 NC_014369 NC_013653 NC_014369 NC_017339	S. aureus S. aureus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.43 1.00
CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 11 stability m NC.011332 NC.015053 NC.010857 NC.015053 NC.010857 NC.017222 NC.004978 NC.004978 NC.001064 NC.010394 NC.010604 NC.013793 NC.005013	P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum E. facaelis E. facaelis E. faceelis	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.63\\ 0.79\\ 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013383 NC_013384 NC_013384 NC_013387 NC_013389 NC_013390 NC_013390 NC_013550 NC_014369 NC_01653 NC_014369 NC_016942 NC_017339 NC_017344	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 : 11 stability n NC.01332 NC.015053 NC.010857 NC.017222 NC.004978 NC.004770 NC.010164 NC.013793 NC.005013 NC.01642 NC.00502	P. marinus P. marinus P. marinus P. marinus P. marinus B. bifidum B. longum B. longum B. longum B. longum B. longum Bifidobacterium sp. M. abscessus M. marinum B. pseudofirmus E. faecalis E. faecalis E. faecalis	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.63\\ 0.79\\ 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013384 NC_013384 NC_013387 NC_013387 NC_013389 NC_013390 NC_013390 NC_013390 NC_013390 NC_013450 NC_014369 NC_016942 NC_017344 NC_017345	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 : 11 stability n NC.01302 NC.015053 NC.015053 NC.010857 NC.017222 NC.004978 NC.004978 NC.004978 NC.010164 NC.010604 NC.010604 NC.0160393 NC.005013 NC.005013	P. marinus P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum B. longum B. longum B. longum Bifidobacterium sp. M. abscessus M. marinum B. pseudofirmus E. faecalis E. faecalis E. faecium E. faecium	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013380 NC_013383 NC_013383 NC_013384 NC_013384 NC_013389 NC_013389 NC_013389 NC_013390 NC_013550 NC_013653 NC_014369 NC_017339 NC_017344 NC_017345	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_008820 NC_009091 NC_009976 : 11 stability n NC_011332 NC_015053 NC_010857 NC_017222 NC_004978 NC_004978 NC_010164 NC_010394 NC_010604 NC_013793 NC_005013 NC_011642 NC_005000 NC_008259	P. marinus P. marinus P. marinus P. marinus P. marinus M. absiliation B. longum B. longum E. fourtient B. pseudofirmus E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013383 NC_013383 NC_013384 NC_013384 NC_013389 NC_013390 NC_013390 NC_013390 NC_013390 NC_013653 NC_013653 NC_014369 NC_017339 NC_017344 NC_017345 NC_017345 NC_017350	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 11 stability n NC.01332 NC.015053 NC.015053 NC.017222 NC.004978 NC.004770 NC.010164 NC.013793 NC.005013 NC.005013 NC.005000 NC.008259 NC.002259 NC.01290	P. marinus P. marinus P. marinus P. marinus P. marinus M. absets23 B. bifidum B. longum B. longum E. facoalis E. faecalis E. faecium E. faecium E. faecium E. faecium	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013383 NC_013384 NC_013384 NC_013387 NC_013389 NC_013390 NC_013390 NC_013390 NC_013550 NC_013653 NC_014369 NC_017344 NC_017345 NC_017350 NC_017350	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 : 11 stability n NC.01322 NC.015053 NC.010857 NC.017222 NC.004978 NC.004770 NC.010164 NC.013793 NC.005013 NC.011642 NC.005003 NC.01642 NC.008259 NC.01290 NC.017962	P. marinus P. marinus P. marinus P. marinus P. marinus P. marinus M. abscessus M. abscessus M. marinum B. pseudofirmus E. faecalis E. faecium E. faecium E. faecium E. faecium E. faecium	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013381 NC_013383 NC_013384 NC_013387 NC_013387 NC_013387 NC_013389 NC_013390 NC_013390 NC_013390 NC_013390 NC_013550 NC_014369 NC_017350 NC_017350 NC_017352 NC_017352 NC_017352 NC_017352	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 : 11 stability n NC.01332 NC.015053 NC.015053 NC.010857 NC.017222 NC.004978 NC.004978 NC.004978 NC.010164 NC.010394 NC.01604 NC.01604 NC.01604 NC.005013 NC.011642 NC.005013 NC.011642 NC.005000 NC.008259 NC.010290 NC.017962 NC.003458	P. marinus P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum B. longum B. longum B. longum Bifidobacterium sp. M. abscessus M. marinum B. pseudofirmus E. faecalis E. faecalis E. faecalis E. faecium E. faecium E. faecium E. faecium E. faecium L. acidophilus L. acidophilus	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Baci	$\begin{array}{c} 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013380 NC_013383 NC_013383 NC_013384 NC_013384 NC_013389 NC_013389 NC_013393 NC_013550 NC_013653 NC_013653 NC_014369 NC_017345 NC_017350 NC_017352 NC_017352 NC_017352 NC_017352 NC_017352 NC_017352 NC_017352	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 : 11 stability m NC.011332 NC.010857 NC.010857 NC.010857 NC.010857 NC.010857 NC.010857 NC.004978 NC.004978 NC.010604 NC.013793 NC.005013 NC.001642 NC.005000 NC.008259 NC.010290 NC.01290 NC.01290 NC.01293 NC.003458 NC.003458 NC.003458	P. marinus P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum E. fourtier S. facealis E. faecalis E. faecalis E. faecium E. faecium E. faecium E. faecium L. acidophilus L. acidophilus	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Baci	$\begin{array}{c} 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013380 NC_013380 NC_013381 NC_013383 NC_013383 NC_013384 NC_013384 NC_013387 NC_013389 NC_013390 NC_013390 NC_013390 NC_013390 NC_013390 NC_013653 NC_014369 NC_017350 NC_017344 NC_017345 NC_017352 NC_018952 NC_018956 NC_018956 NC_018957	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 11 stability n NC.01332 NC.015053 NC.015053 NC.017222 NC.004770 NC.01702 NC.004770 NC.010164 NC.013793 NC.005013 NC.01642 NC.008259 NC.010290 NC.017962 NC.003458 NC.015213 NC.014620	P. marinus P. marinus P. marinus P. marinus P. marinus P. marinus M. absfdum B. longum B. longum E. facosus M. marinum B. pseudofirmus E. faecalis E. faecalis E. faecium E. faecium E. faecium E. faecium L. acidophilus L. acase	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Bacil	$\begin{array}{c} 1.00\\$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013383 NC_013384 NC_013384 NC_013387 NC_013389 NC_013390 NC_013390 NC_013390 NC_013390 NC_013550 NC_013653 NC_014369 NC_017344 NC_017345 NC_017350 NC_017350 NC_017350 NC_017350 NC_017350 NC_017350 NC_017350 NC_018957 NC_018957 NC_018957	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.008820 NC.009091 NC.009976 : 11 stability n NC.01322 NC.015053 NC.010857 NC.017222 NC.004770 NC.010164 NC.010394 NC.004770 NC.010164 NC.013793 NC.005013 NC.011642 NC.008259 NC.010290 NC.017962 NC.015213 NC.015213 NC.016460 NC.016460 NC.016969	P. marinus P. marinus P. marinus P. marinus P. marinus P. marinus P. marinus measure : 0.928523 B. bifidum B. longum B. longum B. longum B. longum B. longum Bifidobacterium sp. M. abscessus M. marinum B. pseudofirmus E. faecalis E. faecalis E. faecium E. faecium	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Bacilli Baci	1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013377 NC_013378 NC_013379 NC_013380 NC_013381 NC_013382 NC_013383 NC_013384 NC_013387 NC_013387 NC_013387 NC_013389 NC_013390 NC_013390 NC_013390 NC_013390 NC_013390 NC_013390 NC_013550 NC_014369 NC_017350 NC_017352 NC_017352 NC_018955 NC_0189557 NC_0189559 NC_0189559 NC_018956	S. aureus S. aureus	Bacilli Bacilli	0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0

PL	NC_018965	S. aureus	Bacilli	1.00	PL	NC_012197	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_018967	S. aureus	Bacilli	0.68	\mathbf{PL}	NC_012198	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_018968	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_012201	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_018972	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_012202	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_018974	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_012203	B. burgdorferi	Spirochaetes	1.00
PL	NC_018976	S. aureus	Bacilli	1.00	PL	NC_{012228}	B. burgdorferi	Spirochaetes	1.00
PL	NC_019007	S. aureus	Bacilli	1.00	$_{\rm PL}$	NC_012231	B. burgdorferi	Spirochaetes	1.00
PL	NC_019008	S. aureus	Bacilli	1.00	PL	NC_012232	B. burgdorferi	Spirochaetes	1.00
PL	NC_019009	S. aureus	Bacilli	1.00	PL	NC_012233	B. burgdorferi	Spirochaetes	1.00
PL	NC_019010	S. aureus	Bacilli	1.00	PL	NC_012236	B. burgdorferi	Spirochaetes	1.00
PL	NC_019148	S. aureus	Bacilli	1.00	PL	NC_012245	B. burgdorferi	Spirochaetes	1.00
PL	NC_019150	S. aureus	Bacilli	1.00	PL	NC_012246	B. burgdorferi	Spirochaetes	1.00
PL	NC_005003	S. epidermiais	Bacilli De silli	1.00		NC_012249	B. burgaorferi B. burgaorferi	Spirochaetes	1.00
PL	NC 005004	S. epidermidis	Bacilli	1.00	PL	NC 012264	B. burgdorferi B. burgdorferi	Spirochaetes	1.00
PL	NC 005006	S. epidermidis	Bacilli	1.00	PL	NC 012494	B. burgdorferi	Spirochaetes	1.00
PL	NC 005566	S epidermidis	Bacilli	1.00	PL	NC 012494	B. burgdorferi B. burgdorferi	Spirochaetes	1.00
PL	NC_006663	S. epidermidis	Bacilli	1.00	PL	NC_012496	B. buradorferi	Spirochaetes	1.00
PL	NC_015219	S. gallolyticus	Bacilli	1.00	PL	NC_012497	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_007171	S. haemolyticus	Bacilli	1.00	$_{\rm PL}$	NC_012498	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_016837	S. infantarius	Bacilli	1.00	\mathbf{PL}	NC_012500	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_007352	S. saprophyticus	Bacilli	1.00	\mathbf{PL}	NC_012502	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_015432	S. saprophyticus	Bacilli	0.50	\mathbf{PL}	NC_012506	B. burgdorferi	Spirochaetes	1.00
PL	NC_016643	S. saprophyticus	Bacilli	0.56	\mathbf{PL}	$NC_{-}012507$	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_013945	S. simulans	Bacilli	1.00	\mathbf{PL}	NC_{012508}	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_005207	S. warneri	Bacilli	1.00	PL	NC_012511	B. burgdorferi	Spirochaetes	1.00
PL	NC_009130	Staphylococcus sp.	Bacilli	1.00	$_{\rm PL}$	NC_017400	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_010938	T. halophilus	Bacilli	1.00	PL	NC_017404	B. burgdorferi	Spirochaetes	1.00
PL	NC_002148	E. rhusiopathiae	Erysipelotrichia	1.00	PL	NC_017405	B. burgdorferi	Spirochaetes	1.00
PL	NC_014634	I. polytropus	Fusobacterna	1.00	PL	NC_017408	B. burgdorferi	Spirochaetes	1.00
PL	NC_007142	C. coli	εproteobacteria	1.00	PL	NC_017411	B. burgdorferi	Spirochaetes	1.00
PL	NC_010858	C. fetus	eproteobacteria	1.00		NC_017412	B. burgaorferi	Spirochaetes	1.00
FL	NC_009715	C. nominis	eproteobacteria	1.00		NC 017413	B. burgdonferi B. burgdonferi	Spirochaetes	1.00
clust	or · 12 stabilit	x = 0.961221			PL	NC 017414	B. burgdorferi B. burgdorferi	Spirochaetes	1.00
PL	NC 008565	B afzelii	Spirochaetes	1.00	PL	NC 017421	B. burgdorferi B. burgdorferi	Spirochaetes	1.00
PL	NC 008566	B. afzelij	Spirochaetes	1.00	PL	NC 018984	B. burgdorferi B. burgdorferi	Spirochaetes	1.00
PL	NC 008567	B. afzelii	Spirochaetes	0.43	PL	NC_018986	B. burgdorferi	Spirochaetes	1.00
PL	NC_008569	B. afzelii	Spirochaetes	1.00	PL	NC_018987	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_011649	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_018988	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_011787	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_018990	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_011788	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_018991	B. burgdorferi	Spirochaetes	1.00
\mathbf{PL}	NC_011792	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	$NC_{-}017775$	B. crocidurae	Spirochaetes	1.00
\mathbf{PL}	NC_011793	B. afzelii	Spirochaetes	1.00	\mathbf{PL}	NC_017798	B. crocidurae	Spirochaetes	1.00
PL	NC_011794	B. afzelii	Spirochaetes	1.00	PL	NC_{017820}	B. crocidurae	Spirochaetes	1.00
PL	NC_017233	B. afzelii	Spirochaetes	1.00	PL	NC_017822	B. crocidurae	Spirochaetes	1.00
\mathbf{PL}	NC_017235	B. afzelii	Spirochaetes	1.00	PL	NC_011254	B. duttonii	Spirochaetes	1.00
\mathbf{PL}	NC_017237	B. afzelii	Spirochaetes	1.00	PL	NC_011854	B. garinii	Spirochaetes	1.00
PL	NC_017239	B. afzelii	Spirochaetes	1.00	PL	NC_011855	B. garınıı	Spirochaetes	1.00
PL	NC_017240	B. afzelii	Spirochaetes	1.00	PL	NC_011857	B. garinii B. sosiaii	Spirochaetes	1.00
PL	NC_015904	B. bissettii	Spirochaetes	1.00		NC_011858	B. garinii B. sosinii	Spirochaetes	1.00
PL	NC_015916	B. Dissettii P. biggettii	Spirochaetes	1.00	PL	NC 011807	Б. yurinii B. aarinii	Spirochaetes	1.00
г L DI	NC 015020	D. Dissettii	Spirochaetes	1.00	PI	NC 011258	B. guinnin B. recurrentis	Spirochaetes	1.00
PL	NC 000950	B huradorferi	Spirochaetes	1.00	PL	NC 011263	B. recurrentis	Spirochaetes	1.00
PL	NC_000955	B. buradorferi	Spirochaetes	1.00	PL	NC_012239	Borrelia sp.	Spirochaetes	1.00
PL	NC_000956	B. burgdorferi	Spirochaetes	0.73	\mathbf{PL}	NC_012247	Borrelia sp.	Spirochaetes	1.00
\mathbf{PL}	NC_000957	B. burgdorferi	Spirochaetes	1.00	PL	NC_012260	Borrelia sp.	Spirochaetes	1.00
\mathbf{PL}	NC_001851	B. burgdorferi	Spirochaetes	1.00	\mathbf{PL}	NC_012176	B. spielmanii	Spirochaetes	1.00
\mathbf{PL}	NC_{001852}	B. burgdorferi	Spirochaetes	1.00	\mathbf{PL}	NC_012183	B. spielmanii	Spirochaetes	1.00
\mathbf{PL}	NC_001853	B. burgdorferi	Spirochaetes	1.00	PL	NC_012190	B. spielmanii	Spirochaetes	1.00
PL	NC_001855	B. burgdorferi	Spirochaetes	1.00	PL	NC_012204	B. valaisiana	Spirochaetes	1.00
PL	NC_001856	B. burgdorferi	Spirochaetes	1.00	- · ·				
PL	NC_011736	B. burgdorferi	Spirochaetes	1.00	cluster	: 13 stability	measure : 0.837212		1 0 0
PL	NC_011778	B. burgdorferi	Spirochaetes	1.00	CHR	NC_015690	P. mucilaainosus	Bacilli	1.00
PL DI	NC_011779			1 1	UHR	NC 016005	P mag :1:	D a	
PI.	NC 011790	B huradorferi	Spirochaetes	1.00	СПВ	NC_016935	P. mucilaginosus P. mucilaginosus	Bacilli	1.00
	NC_011780 NC_011781	B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00	CHR	NC_016935 NC_017672 NC_012914	P. mucilaginosus P. mucilaginosus Paenihacillus sp	Bacilli Bacilli Bacilli	1.00
PL	NC_011780 NC_011781 NC_011782	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57	CHR CHR CHR	NC_016935 NC_017672 NC_012914 NC_013406	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp.	Bacilli Bacilli Bacilli Bacilli	1.00 1.00 0.77 1.00
PL PL	NC_011780 NC_011781 NC_011782 NC_011785	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36	CHR CHR CHR	NC_016935 NC_017672 NC_012914 NC_013406	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp.	Bacilli Bacilli Bacilli Bacilli	1.00 1.00 0.77 1.00
PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011849	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 0.57 0.36 1.00	CHR CHR CHR CHR	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528	Bacilli Bacilli Bacilli Bacilli	1.00 1.00 0.77 1.00
PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011849 NC_011864	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	CHR CHR CHR CHR <u>cluster</u> PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor	Bacilli Bacilli Bacilli Bacilli Actinobacteridae	1.00 1.00 0.77 1.00
PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011849 NC_011864 NC_011868	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$ 1.00 \\ 1.00 \\ 1.00 \\ 0.57 \\ 0.36 \\ 1.00$	CHR CHR CHR CHR PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae	1.00 1.00 0.77 1.00 0.87 1.00
PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011869 NC_011864 NC_011868 NC_011870	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$ \begin{array}{c} 1.00\\ 1.00\\ 0.57\\ 0.36\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array} $	CHR CHR CHR CHR PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009467	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae aproteobacteria	1.00 1.00 0.77 1.00 0.87 1.00 0.37
PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011864 NC_011864 NC_011868 NC_011870 NC_011872	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR CHR PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_003904 NC_009467 NC_009469	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria	1.00 1.00 0.77 1.00 0.87 1.00 0.37 0.37
PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011849 NC_011864 NC_011868 NC_011870 NC_011870 NC_011872 NC_011874	B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.57\\ 0.36\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	CHR CHR CHR PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_005952 NC_009467 NC_009469 NC_015178	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria	1.00 1.00 0.77 1.00 0.87 1.00 0.37 0.37 0.29
PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011849 NC_011868 NC_011868 NC_011870 NC_011872 NC_011874 NC_011965	B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$\begin{array}{c} 1.00\\ 1.00\\ 0.57\\ 0.36\\ 1.00\\$	CHR CHR CHR PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009467 NC_009469 NC_015178 NC_003064	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum A. fabrum	Bacilli Bacilli Bacilli Bacilli Actinobacteridae αproteobacteria αproteobacteria αproteobacteria αproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.77\\ 1.00\\ 0.87\\ 1.00\\ 0.37\\ 0.37\\ 0.29\\ 0.81\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011865 NC_011866 NC_011868 NC_011870 NC_011872 NC_011874 NC_011965 NC_011972	 b. burgdorferi b. burgdorferi b. burgdorferi b. burgdorferi b. burgdorferi b. burgdorferi burgdorferi 	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.57\\ 0.36\\ 1.00\\$	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009467 NC_009467 NC_003065 NC_003065	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum A. fabrum A. fabrum	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria αproteobacteria αproteobacteria	1.00 1.00 0.77 1.00 0.87 1.00 0.37 0.37 0.29 0.81 1.00
PL PL PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011865 NC_011864 NC_011868 NC_011870 NC_011870 NC_011872 NC_011965 NC_011972 NC_012105	B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.57\\ 0.36\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_003904 NC_009469 NC_009469 NC_003064 NC_003065 NC_011987	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum A. fabrum A. fabrum A. radiobacter	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.77\\ 1.00\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011868 NC_011870 NC_011872 NC_011872 NC_011874 NC_011965 NC_011972 NC_012105 NC_012105	B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_003904 NC_009469 NC_009469 NC_003064 NC_003065 NC_011987 NC_011990	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum A. fabrum A. fabrum A. radiobacter A. radiobacter	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria	1.00 1.00 0.77 1.00 0.87 1.00 0.37 0.37 0.37 0.29 0.81 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011864 NC_011864 NC_011868 NC_011870 NC_011872 NC_011872 NC_011972 NC_012105 NC_012105 NC_012162	B. burgdorferi B. burgdorferi	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_003904 NC_009469 NC_009469 NC_003064 NC_003064 NC_003065 NC_011987 NC_011990 NC_002575	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. chir	Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria	1.00 1.00 0.77 1.00 0.37 0.37 0.29 0.81 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011864 NC_011864 NC_011868 NC_011868 NC_011872 NC_011874 NC_011965 NC_011972 NC_011965 NC_012156 NC_012162 NC_012162 NC_012163	 burgdorferi 	Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009467 NC_009467 NC_003065 NC_003065 NC_011987 NC_0010575 NC_011990 NC_002575 NC_010841	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. thizogenes	Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.77\\ 1.00\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011868 NC_011870 NC_011870 NC_011870 NC_011874 NC_011965 NC_012165 NC_012162 NC_012163 NC_01269	 b. burgdorferi b. burgdorferi burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_009467 NC_009467 NC_009469 NC_015178 NC_003064 NC_003065 NC_011987 NC_011990 NC_002575 NC_010841 NC_002147	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. multivorum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. thumefaciens A tumefaciens	Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.77\\ 1.00\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011868 NC_011870 NC_011870 NC_011870 NC_011870 NC_011875 NC_011965 NC_012165 NC_012162 NC_012163 NC_012167 NC_012168 NC_012168	 burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009467 NC_009469 NC_015178 NC_003065 NC_011987 NC_011990 NC_002575 NC_010841 NC_002147 NC_0022077	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. tumefaciens A. tumefaciens A. tumefaciens	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.77\\ 1.00\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011868 NC_011868 NC_011870 NC_011872 NC_011872 NC_011874 NC_011972 NC_012105 NC_012105 NC_012163 NC_012167 NC_012167 NC_012171	 burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_003904 NC_009469 NC_009469 NC_003065 NC_011987 NC_003065 NC_011987 NC_002575 NC_010841 NC_002147 NC_002377 NC_010929 NC_011981	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. multivorum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. radiobacter A. rhizogenes A. thunefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens	Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.77\\ 1.00\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011866 NC_011870 NC_011872 NC_011872 NC_011872 NC_011874 NC_011965 NC_012165 NC_012162 NC_012162 NC_012163 NC_012168 NC_012171 NC_012182	 burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_003904 NC_003064 NC_003064 NC_003065 NC_011987 NC_001987 NC_002175 NC_011990 NC_002377 NC_002377 NC_010929 NC_011981 NC_011982	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. multivorum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. thumefaciens A. tumefaciens A. tumefaciens A. vitis	Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	1.00 1.00 0.077 1.00 0.87 1.00 0.37 0.37 0.37 0.37 0.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011868 NC_011868 NC_011872 NC_011872 NC_011874 NC_011965 NC_012165 NC_012166 NC_012162 NC_012167 NC_012168 NC_012170 NC_012171 NC_012184	 burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_009467 NC_009467 NC_009467 NC_009469 NC_003065 NC_011987 NC_011987 NC_002575 NC_010841 NC_002147 NC_002377 NC_010929 NC_011984	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. Paenibacillus sp. S. coelicolor S. violaceusniger A. cryptum A. cryptum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. thunefaciens A. tumefaciens A. tumefaciens A. vitis A. vitis	Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae αproteobacteria	0.87 1.00 0.77 1.00 0.77 1.00 0.77 1.00 0.37 0.37 0.29 0.81 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011868 NC_011870 NC_011870 NC_011870 NC_011872 NC_011972 NC_012165 NC_012163 NC_012163 NC_012163 NC_012167 NC_012168 NC_012170 NC_012171 NC_012182 NC_012184 NC_012186	 b. burgdorferi b. burgdorferi b. burgdorferi b. burgdorferi b. burgdorferi b. burgdorferi burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009467 NC_009467 NC_009469 NC_015975 NC_010987 NC_010981 NC_002575 NC_010841 NC_002147 NC_002147 NC_00299 NC_011981 NC_011984 NC_011984	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. Paenibacillus sp. S. coelicolor S. violaceusniger A. cryptum A. cryptum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. vitis A. vitis A. vitis A. vitis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae Aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria	1.00 1.00 1.00 0.77 1.00 0.77 1.00 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.37 0.29 0.81 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011868 NC_011870 NC_011870 NC_011870 NC_011870 NC_011972 NC_012105 NC_012162 NC_012163 NC_012167 NC_012167 NC_012170 NC_012171 NC_012184 NC_012186 NC_012186 NC_012189	 b. burgdorferi burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009467 NC_009469 NC_015178 NC_003065 NC_011987 NC_011980 NC_002575 NC_010841 NC_002377 NC_010841 NC_002377 NC_010929 NC_011981 NC_011982 NC_011984 NC_011986 NC_011991	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. rhizogenes A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. vitis A. vitis A. vitis A. vitis A. vitis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae Actinobacteridae aproteobacteria	1.00 1.00 0.77 1.00 0.77 1.00 0.37 0.29 0.81 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011868 NC_011868 NC_011870 NC_011870 NC_011872 NC_011972 NC_012105 NC_012105 NC_012163 NC_012163 NC_012163 NC_012167 NC_012163 NC_012171 NC_012182 NC_012184 NC_012189 NC_012189 NC_012192	 burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009469 NC_015178 NC_003065 NC_011987 NC_003065 NC_011987 NC_002575 NC_010841 NC_002147 NC_002377 NC_010921 NC_011981 NC_011981 NC_011984 NC_011984 NC_011991 NC_015184	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. cryptum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. vitis A. vitis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae aproteobacteria	1.00 1.00 0.077 1.00 0.77 1.00 0.37 0.29 0.81 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011849 NC_011864 NC_011868 NC_011868 NC_011870 NC_011874 NC_011874 NC_011965 NC_012165 NC_012162 NC_012162 NC_012163 NC_012168 NC_012168 NC_012170 NC_012182 NC_012184 NC_012184 NC_012189 NC_012192	 burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_016935 NC_017672 NC_012914 NC_013406 : 14 stability NC_003904 NC_015952 NC_009469 NC_015178 NC_003065 NC_011987 NC_002147 NC_002147 NC_002377 NC_010929 NC_011981 NC_011982 NC_011984 NC_011984 NC_011984 NC_011984 NC_011984 NC_015184 NC_016619	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. measure : 0.668528 S. coelicolor S. violaceusniger A. cryptum A. multivorum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. thuefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. vitis A. vitis A. vitis A. vitis A. vitis A. vitis A. vitis A. vitis A. spobacterium sp. A. brasilense	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae αproteobacteria	1.00 1.00 1.00 0.77 1.00 0.77 1.00 0.37 0.37 0.37 0.39 0.81 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_011780 NC_011781 NC_011782 NC_011785 NC_011785 NC_011864 NC_011868 NC_011868 NC_011872 NC_011872 NC_011874 NC_011965 NC_012165 NC_012167 NC_012168 NC_012167 NC_012170 NC_012170 NC_012184 NC_012184 NC_012186 NC_012189 NC_012195 NC_012195 NC_012195	 burgdorferi 	Spirochaetes Spirochaetes	1.00 1.00 1.00 1.00 0.57 0.36 1.00	CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	$\begin{array}{c} {\rm NC}_{-}016935 \\ {\rm NC}_{-}017672 \\ {\rm NC}_{-}012914 \\ {\rm NC}_{-}013406 \\ \hline \\ {\rm : 14 \ stability} \\ {\rm NC}_{-}003064 \\ {\rm NC}_{-}003064 \\ {\rm NC}_{-}003064 \\ {\rm NC}_{-}003065 \\ {\rm NC}_{-}003064 \\ {\rm NC}_{-}003065 \\ {\rm NC}_{-}011987 \\ {\rm NC}_{-}002575 \\ {\rm NC}_{-}010841 \\ {\rm NC}_{-}002575 \\ {\rm NC}_{-}010841 \\ {\rm NC}_{-}002477 \\ {\rm NC}_{-}002377 \\ {\rm NC}_{-}010929 \\ {\rm NC}_{-}011981 \\ {\rm NC}_{-}011984 \\ {\rm NC}_{-}011984 \\ {\rm NC}_{-}015184 \\ {\rm NC}_{-}016619 \\ {\rm NC}_{-}012847 \\ \end{array}$	P. mucilaginosus P. mucilaginosus Paenibacillus sp. Paenibacillus sp. Paenibacillus sp. S. coelicolor S. violaceusniger A. cryptum A. cryptum A. cryptum A. fabrum A. fabrum A. fabrum A. fabrum A. fabrum A. radiobacter A. radiobacter A. rhizogenes A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. tumefaciens A. vitis A. vitis A. vitis A. vitis A. vitis A. vitis A. vitis A. vitis A. vitis A. sitis A.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Actinobacteridae aproteobacteria	1.00 1.00 1.00 0.77 1.00 0.77 1.00 0.37 0.00 1.00 1.00 1.00 1.00 1.00 0.00 1.00 0.95 0.00

PL	NC 010590	Diadian		1 00	DI	NC 002220	T	D:11:	0.00
1.12	NC 009475	Bradurhizohium sp	aproteobacteria	1.00	PL	NC 016970	L. curviege	Bacilli	1.00
PL I	NC 010333	Caulobacter sp.	aproteobacteria	0.32	PL	NC 016971	L. garviege	Bacilli	1.00
PL	NC 008242	Chalatinorane sp.	aproteobacteria	1.00	PI	NC 011830	L. gasseri	Bacilli	0.64
PL	NC 008242	Chelativorans sp.	aproteobacteria	1.00	PL	NC 001379	L. gasseri	Bacilli	0.04
DI	NC 008243	Chelatiworana sp.	aproteobacteria	1.00		NC 002102	L. netveticus	Dacilli	1.00
PL	NC 000055	D shihae	aproteobacteria	0.50	PI	NC 014386	L. helveticus	Bacilli	0.64
DI	NC 009957	D. shibae	aproteobacteria	1.00	PL	NC 017468	L. helveticus	Bacilli	0.04
PL	NC 009958	D. shibae	aproteobacteria	1.00	PL	NC 015603	L. kefiranofaciens	Bacilli	1 00
PL	NC 016021	G rulinus	aproteobacteria	0.00	PL	NC 014131	L. kimchii	Bacilli	1.00
PL	NC 016029	G rulinus	aproteobacteria	0.62	PL	NC 014132	L kimchii	Bacilli	1 00
PL	NC 014918	M. ciceri	aproteobacteria	1.00	PL	NC 014133	L. kimchii	Bacilli	1.00
PL	NC_002679	M. loti	aproteobacteria	0.21	PL	NC_014134	L. kimchii	Bacilli	1.00
PL	NC 002682	M. loti	aproteobacteria	1.00	PL	NC 000906	L. lactis	Bacilli	1.00
PL	NC 007959	N hamburaensis	aproteobacteria	0.73	PL	NC 001949	L lactis	Bacilli	1.00
PL	NC 007960	N. hamburgensis	aproteobacteria	1.00	PL	NC 002137	L. lactis	Bacilli	1.00
PL	NC_007961	N. hamburgensis	aproteobacteria	1.00	PL	NC_002138	L. lactis	Bacilli	1.00
PL	NZ_AGFM	N. pentaromativorans	αproteobacteria	0.44	PL	NC_002150	L. lactis	Bacilli	1.00
	01000122	1			\mathbf{PL}	NC_002193	L. lactis	Bacilli	1.00
PL	NC_009669	O. anthropi	α proteobacteria	0.50	PL	NC_002502	L. lactis	Bacilli	nan
PL	NC_009670	O. anthropi	α proteobacteria	1.00	\mathbf{PL}	NC_002798	L. lactis	Bacilli	1.00
PL	NC_009671	O. anthropi	α proteobacteria	1.00	\mathbf{PL}	NC_003101	L. lactis	Bacilli	1.00
PL	NC_015685	O. carboxidovorans	α proteobacteria	1.00	\mathbf{PL}	NC_004163	L. lactis	Bacilli	1.00
PL	NC_015689	O. carboxidovorans	α proteobacteria	1.00	PL	NC_004164	L. lactis	Bacilli	1.00
PL	NC_017536	O. carboxidovorans	α proteobacteria	1.00	PL	NC_004652	L. lactis	Bacilli	1.00
PL	NC_017539	O. carboxidovorans	α proteobacteria	1.00	\mathbf{PL}	NC_004653	L. lactis	Bacilli	1.00
PL	NC_015258	P. gilvum	α proteobacteria	1.00	PL	NC_004847	L. lactis	Bacilli	1.00
PL	NC_014035	R. capsulatus	α proteobacteria	1.00	\mathbf{PL}	NC_004955	L. lactis	Bacilli	1.00
PL	NC_008386	R. denitrificans	α proteobacteria	0.62	PL	NC_004959	L. lactis	Bacilli	nan
PL	NC_004041	R. etli	α proteobacteria	1.00	\mathbf{PL}	NC_004960	L. lactis	Bacilli	1.00
PL	NC_007762	R. etli	α proteobacteria	1.00	PL	NC_004966	L. lactis	Bacilli	1.00
PL	NC_007763	R. etli	α proteobacteria	1.00	PL	NC_007191	L. lactis	Bacilli	1.00
PL	NC_007764	R. etli	α proteobacteria	nan	PL	NC_008436	L. lactis	Bacilli	1.00
PL	NC_007765	R. etli	α proteobacteria	1.00	PL	NC_008503	L. lactis	Bacilli	1.00
PL	NC_007766	R. etli	α proteobacteria	1.00	PL	NC_008504	L. lactis	Bacilli	1.00
PL	NC_010996	R. eth	α proteobacteria	1.00	PL	NC_008505	L. lactis	Bacilli	1.00
PL	NC 0100997	R. etti P. otli	αproteobacteria	1.00	PL	NC 008507	L. lactis	Bacilli	1.00
PL DI	NC 008278	R. logumin o gamum	aproteobacteria	0.71		NC 008504	L. lactic	Bacilli	1.00
PL	NC 008381	R leguminosarum	aproteobacteria	0.91	PL	NC 009137	L. lactis	Bacilli	1.00
PL	NC 008382	R leguminosarum	aproteobacteria	1.00	PL	NC 009751	L lactis	Bacilli	1 00
PL	NC_008383	R. leguminosarum	α proteobacteria	1.00	\mathbf{PL}	NC_010901	L. lactis	Bacilli	0.43
PL	NC_008384	R. leguminosarum	α proteobacteria	1.00	\mathbf{PL}	NC_015860	L. lactis	Bacilli	1.00
PL	NC_011366	R. leguminosarum	α proteobacteria	1.00	\mathbf{PL}	NC_015861	L. lactis	Bacilli	1.00
PL	NC_011370	$R. \ leguminos arum$	α proteobacteria	1.00	PL	NC_015862	L. lactis	Bacilli	1.00
PL	NC_012848	R. leguminosarum	α proteobacteria	0.68	\mathbf{PL}	NC_015863	L. lactis	Bacilli	1.00
PL	NC_012853	R. leguminosarum	α proteobacteria	0.77	PL	NC_015900	L. lactis	Bacilli	1.00
PL	NC_012854	R. leguminosarum	α proteobacteria	1.00	PL	NC_015901	L. lactis	Bacilli	1.00
PL	NC_012858	R. leguminosarum	αproteobacteria	0.77	PL	NC_015912	L. lactis	Bacilli	1.00
PL	NC 004574	R. litoralis	αproteobacteria	1.00	PL	NC 017478	L. lactis	Bacilli	1.00
PL	NC 007488	Ruegeriu sp.	aproteobacteria	1.00	PL	NC 017478	L. lactis	Bacilli	1.00
PL	NC 011962	R sphaeroides	aproteobacteria	1.00	PL	NC 017484	L. lactis	Bacilli	1.00
PL	NZ_CM001164	R. sphaeroides	aproteobacteria	0.94	PL	NC_017485	L. lactis	Bacilli	1.00
PL	NC_000914	S. fredii	α proteobacteria	0.44	\mathbf{PL}	NC_017487	L. lactis	Bacilli	1.00
PL	NC_015742	S. fredii	α proteobacteria	1.00	\mathbf{PL}	NC_017489	L. lactis	Bacilli	1.00
PL		J		0.43	PL	NC 017493	L. lactis		
	NC_016814	S. fredii	α proteobacteria					Bacilli	1.00
PL	NC_016814 NC_009621	S. fredii S. medicae	lphaproteobacteria $lpha$ proteobacteria	0.87	\mathbf{PL}	NC_017498	L. lactis	Bacilli Bacilli	$1.00 \\ 1.00$
PL PL	NC_016814 NC_009621 NC_009622	S. fredii S. medicae S. medicae	lphaproteobacteria $lpha$ proteobacteria $lpha$ proteobacteria	0.87 0.67	$_{\rm PL}^{\rm PL}$	NC_017498 NC_017500	L. lactis L. lactis	Bacilli Bacilli Bacilli	$1.00 \\ 1.00 \\ 1.00$
PL PL PL	NC_016814 NC_009621 NC_009622 NC_003037	S. fredii S. medicae S. medicae S. meliloti	lphaproteobacteria lphaproteobacteria lphaproteobacteria lphaproteobacteria	0.87 0.67 nan	PL PL PL	NC_017498 NC_017500 NC_008496	L. lactis L. lactis L. mesenteroides	Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 0.73
PL PL PL PL	NC_016814 NC_009621 NC_009622 NC_003037 NC_010865	S. fredii S. medicae S. medicae S. meliloti S. meliloti	α proteobacteria α proteobacteria α proteobacteria α proteobacteria α proteobacteria	0.87 0.67 nan 1.00	PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820	L. lactis L. lactis L. mesenteroides L. mesenteroides	Bacilli Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 0.73 1.00
PL PL PL PL PL	NC_016814 NC_009621 NC_009622 NC_003037 NC_010865 NC_013545	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti	αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.87 0.67 nan 1.00 0.00	PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 0.73 1.00 0.73 1.00
PL PL PL PL PL RECE*	NC_016814 NC_009621 NC_009622 NC_003037 NC_010865 NC_013545 NC_015591 NC_015591	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti	αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.87 0.67 nan 1.00 0.00 0.33	PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	1.00 1.00 0.73 1.00 0.73 1.00 1.00
PL PL PL PL RECE* PL DL	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015591 NC.015597 NC.015397	S. fredii S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti	αproteobacteria	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36	PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_003894	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$ \begin{array}{c} 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 1.00\\ 1.00\\ 0.26\end{array} $
PL PL PL PL PL PL PL PL PL	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015597 NC.017324 NC.017324	S. fredii S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti	 αproteobacteria 	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.32	PL PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_003894 NC_004944	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	1.00 1.00 0.73 1.00 0.73 1.00 1.00 0.36 1.00
PL PL PL PL PL PL PL PL PL	NC_0016814 NC_009622 NC_009622 NC_001865 NC_010865 NC_013545 NC_015591 NC_015597 NC_017324 NC_017327 NC_01862	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti	αproteobacteria	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00	PL PL PL PL PL PL PL PL PJ	NC-017498 NC-017500 NC-008496 NC-016820 NC-016821 NC-016828 NC-003894 NC-004944 NC-006278 NC 006377	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 1.00\\ 0.36\\ 1.00\\ 0.43\end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_006814 NC_009621 NC_009622 NC_003037 NC_010865 NC_013545 NC_015597 NC_015597 NC_017324 NC_017327 NC_018682 NC_018683	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti	 αproteobacteria 	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.00	PL PL PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_003894 NC_004944 NC_006278 NC_006377 NC_011101	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 0.43\\ 0.43\end{array}$
PL PL PL PL PL PL PL PL PL PL PL	NC_0016814 NC_009621 NC_009622 NC_003037 NC_010865 NC_013545 NC_015597 NC_015597 NC_017324 NC_017327 NC_018682 NC_018683 NC_009717	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti	αproteobacteria	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.00 1.00	PL PL PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_003894 NC_004944 NC_006278 NC_006278 NC_006377 NC_014558	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 0.43\\ 1.00\end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL	NC.016814 NC.009621 NC.009622 NC.003037 NC.013545 NC.013545 NC.015591 NC.017324 NC.017327 NC.018682 NC.018683 NC.009717 NC.006824	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum	 αproteobacteria 	$\begin{array}{c} 0.87\\ 0.67\\ nan\\ 1.00\\ 0.00\\ \hline 0.33\\ 1.00\\ 0.36\\ 0.33\\ 1.00\\ 0.00\\ 1.00\\ 0.23\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_00828 NC_003894 NC_006278 NC_006277 NC_011101 NC_014558 NC_010621	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 0.43\\ 1.00\\ 0.36\end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_0016814 NC_009622 NC_009622 NC_001865 NC_013545 NC_0135591 NC_015597 NC_017324 NC_017327 NC_018682 NC_018683 NC_008683 NC_008683 NC_008760	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans	αproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 0.87\\ 0.67\\ nan\\ 1.00\\ 0.00\\ 0.33\\ 1.00\\ 0.36\\ 0.33\\ 1.00\\ 0.00\\ 1.00\\ 0.23\\ 0.33\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_003894 NC_006278 NC_006278 NC_006377 NC_011101 NC_014558 NC_016621 NC_015698	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_006814 NC_009622 NC_009622 NC_001865 NC_013545 NC_015591 NC_015597 NC_017324 NC_017324 NC_017327 NC_018682 NC_018683 NC_008717 NC_00824 NC_008241	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha	αproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	0.87 0.67 nan 1.00 0.33 1.00 0.33 1.00 0.33 1.00 0.00 1.00 0.23 0.33 0.33	PL PL PL PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_00828 NC_003894 NC_004944 NC_006278 NC_006377 NC_011101 NC_014558 NC_006621 NC_015698 NC_015701	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_006814 NC_009622 NC_009622 NC_003037 NC_010865 NC_013545 NC_015591 NC_015597 NC_017324 NC_017327 NC_018682 NC_018683 NC_008717 NC_006824 NC_008741 NC_005241 NC_008739	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti A. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei	 αproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria 	$\begin{array}{c} 0.87\\ 0.67\\ nan\\ 1.00\\ 0.00\\ \hline 0.33\\ 1.00\\ 0.36\\ 0.33\\ 1.00\\ 0.00\\ 1.00\\ 0.23\\ 0.33\\ 0.32\\ 0.25\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL PL	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_003894 NC_004944 NC_004944 NC_006278 NC_004974 NC_014558 NC_014558 NC_016598 NC_015200	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. reuteri L. rhamnosus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_016814 NC_009621 NC_009622 NC_003037 NC_010865 NC_013545 NC_015591 NC_015597 NC_017324 NC_017324 NC_017327 NC_018683 NC_009717 NC_006824 NC_008760 NC_008739	S. fredii S. medicae S. medicat S. meliloti S. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei	$\begin{array}{c} \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	$\begin{array}{c} 0.87\\ 0.67\\ nan\\ 1.00\\ 0.00\\ \hline 0.33\\ 1.00\\ 0.36\\ 0.33\\ 1.00\\ 0.00\\ 1.00\\ 0.23\\ 0.33\\ 0.32\\ 0.25\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL PL PL	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.004944 NC.004944 NC.006278 NC.006278 NC.014558 NC.016521 NC.016598 NC.015698 NC.013200 NC.004942	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. rehamnosus L. sakei	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.013545 NC.013545 NC.015591 NC.017324 NC.017327 NC.017327 NC.018682 NC.018683 NC.008717 NC.008624 NC.008760 NC.008739	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515	α proteobacteria α proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria	$\begin{array}{c} 0.87\\ 0.67\\ nan\\ 1.00\\ 0.00\\ \hline 0.33\\ 1.00\\ 0.33\\ 1.00\\ 0.33\\ 1.00\\ 0.00\\ 1.00\\ 0.23\\ 0.33\\ 0.32\\ 0.25\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.00828 NC.003894 NC.006278 NC.006278 NC.006277 NC.011101 NC.016528 NC.015698 NC.015701 NC.013200 NC.004942 NC.006529	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. reuteri L. sakei L. salivarius	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.001865 NC.013545 NC.0135591 NC.017324 NC.017327 NC.017327 NC.018682 NC.018683 NC.009717 NC.006824 NC.008760 NC.008760 NC.008760 NC.008739 I5 stability meaa NC.014333	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus	αproteobacteria βproteobacteria βproteobacteri	0.87 0.67 nan 1.00 0.33 1.00 0.33 1.00 0.33 0.33 0.32 0.25 0.73	PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.00828 NC.003894 NC.006278 NC.006278 NC.006277 NC.011101 NC.016528 NC.015698 NC.015701 NC.013200 NC.006529 NC.006529 NC.006530	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 0.50\\ 0.50\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015591 NC.015597 NC.017324 NC.017327 NC.018682 NC.018683 NC.008717 NC.008683 NC.008760 NC.008760 Stability mear NC.014333 NC.010853	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans	αproteobacteria βproteobacteria βproteobacteri	0.87 0.67 nan 1.00 0.33 1.00 0.33 1.00 0.33 0.33 0.32 0.25	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.00828 NC.003894 NC.006278 NC.006278 NC.006278 NC.01101 NC.014558 NC.015698 NC.015698 NC.015701 NC.015299 NC.006529 NC.006529 NC.006530 NC.017479	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. reuteri L. sakei L. salivarius L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.43\\ 1.00\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 0.50\\ 1.00\\ 0.47\\ 0.50\\ 1.00\\ 0.47\\ 0.50\\ 1.00\\ 0.47\\ 0.50\\ 1.00\\ 0.47\\ 0.50\\ 0.00\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015591 NC.015597 NC.017324 NC.017324 NC.017327 NC.018683 NC.009717 NC.008683 NC.008741 NC.0085241 NC.0085241 NC.004333 NC.014333 NC.010853 NC.01553 NC.005510	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis F. faecalis	αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.23 0.23 0.25	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_00394 NC_004944 NC_004944 NC_004944 NC_004944 NC_004578 NC_014558 NC_01651 NC_014558 NC_01659 NC_006529 NC_006530 NC_017479 NC_017480	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. relateri L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius L. salivarius L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.43\\ 1.00\\ 0.33\\ 0.03\\ 0.030\\ 0.30\\ 0.50\\ 1.00\\ 0.47\\ 0.50\\ 1.00\\ 0.00$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015591 NC.017324 NC.017327 NC.018682 NC.018683 NC.008717 NC.008824 NC.008760 NC.008541 NC.008739 I5 stability mear NC.014333 NC.010853 NC.010853 NC.010853 NC.010853	S. fredii S. medicae S. medicae S. meliloti S. meliloti M. aquacolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis	αproteobacteria βproteobacteria βeriebacteria βeriebacteria βeriebacteria βeriebacteria βeriebacteria	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.00 0.23 0.23 0.25 0.73 1.00 0.36 0.36	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_003894 NC_004944 NC_004944 NC_004944 NC_004944 NC_004944 NC_014558 NC_010621 NC_014558 NC_016529 NC_006529 NC_007479 NC_017479 NC_017480 NC_015979 NC_004942	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. salivarius L. salivarius L. salivarius L. salivarius L. salivarius L. salivarius L. salivarius L. salivarius L. sanifranciscensis	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.47\\ 0.50\\ 1.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.72\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013845 NC.015597 NC.017324 NC.017327 NC.018682 NC.017327 NC.018683 NC.009717 NC.006824 NC.008739 IS stability mean NC.014333 NC.010853 NC.010853 NC.005010 NC.013533 NC.014508 NC.015319	S. fredii S. medicae S. medicae S. meliloti S. meliloti M. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. ccreus B. caegulans E. faecalis E. faecalis E. faecalis L. amulonorus	αproteobacteria βproteobacteria βeriteobacteria βeriteobacteria βeriteobacteria βeriteobacteria βeriteobacteria βeriteobacteria βeriteobacteria βeriteobacteria βeriteobacteria βeriteobacteri	0.87 0.67 nan 1.00 0.33 1.00 0.36 0.33 1.00 0.036 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 0.00 0.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.016820 NC.016820 NC.016821 NC.016828 NC.004944 NC.004944 NC.006278 NC.004944 NC.006278 NC.01458 NC.016621 NC.016698 NC.015698 NC.015479 NC.017479 NC.017480 NC.015979 NC.017480 NC.01608	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. sanfranciscensis P. acidilactici	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.30\\ 0.50\\ 1.00\\ 0.47\\ 1.00\\ 0.50\\ 1.00\\ 0.70\\ 0.00\\ 0.70\\ 0.47\end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.00365 NC.013545 NC.013545 NC.015597 NC.017324 NC.017327 NC.017327 NC.018682 NC.018682 NC.008717 NC.006824 NC.008760 NC.008760 NC.008739 I5 stability mean NC.014333 NC.014533 NC.005010 NC.013533 NC.00510 NC.013533 NC.0015319 NC.008498	S. fredii S. medicae S. medicae S. meliloti S. meliloti M. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. coreus B. coagulans E. faecalis E. faecalis E. faecalis L. anylovorus L. brevis	αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli <th>0.87 0.67 nan 1.00 0.33 1.00 0.36 0.33 1.00 0.00 1.00 0.23 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 nan 0.27</th> <th>PL PL PL PL PL PL PL PL PL PL PL PL PL P</th> <th>NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.006278 NC.006278 NC.006278 NC.010621 NC.015698 NC.015698 NC.015698 NC.015698 NC.004942 NC.004942 NC.006529 NC.004530 NC.007479 NC.017480 NC.017480 NC.017480 NC.017608 NC.016636</th> <th>L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius</th> <th>Bacilli Bacilli</th> <th>$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.33\\ 1.00\\ 0.43\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.47\\ 1.00\\ 0.47\\$</th>	0.87 0.67 nan 1.00 0.33 1.00 0.36 0.33 1.00 0.00 1.00 0.23 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 nan 0.27	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.006278 NC.006278 NC.006278 NC.010621 NC.015698 NC.015698 NC.015698 NC.015698 NC.004942 NC.004942 NC.006529 NC.004530 NC.007479 NC.017480 NC.017480 NC.017480 NC.017608 NC.016636	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.33\\ 1.00\\ 0.43\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.00\\ 0.00\\ 0.00\\ 0.00\\ 0.47\\ 1.00\\ 0.47\\$
PL PL PL PL PL PL PL PL PL PL	NC.016814 NC.009621 NC.009622 NC.00365 NC.013865 NC.0135591 NC.015597 NC.017324 NC.017327 NC.018682 NC.018683 NC.009717 NC.006824 NC.008760 NC.008760 NC.008739 I5 stability meai NC.014333 NC.010853 NC.005010 NC.013533 NC.014508 NC.008499	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis L. amylovorus L. brevis	αproteobacteria βproteobacteria βproteobacteria βproteobacteria γproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.87 0.67 nan 1.00 0.33 1.00 0.33 1.00 0.33 1.00 0.00 1.00 0.23 0.33 0.25 0.73 1.00 0.23 0.25	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.006278 NC.006278 NC.006278 NC.010621 NC.016698 NC.015698 NC.015698 NC.015699 NC.004529 NC.004529 NC.004529 NC.004530 NC.017479 NC.017479 NC.015979 NC.015979 NC.004832 NC.016608 NC.016636 NC.017017	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. rateri L. sakivarius L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.33\\ 1.00\\ 0.33\\ 1.00\\ 0.43\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.30\\ 0.47\\ 1.00\\ 0.00\\ 0.70\\ 0.00\\ 0.71\\ 1.00\\ 0.67\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015591 NC.015597 NC.017324 NC.017324 NC.017327 NC.018682 NC.018683 NC.008717 NC.006824 NC.006824 NC.006824 NC.0008241 NC.008739 I5 stability mear NC.014333 NC.014533 NC.014533 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.014508 NC.012550	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis L. brevis L. brevis L. brevis	αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.23 0.23 0.25 0.73 1.00 0.36 0.30 0.36 0.00 nan 0.27 0.00 na57	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_017498 NC_017500 NC_008496 NC_016820 NC_016821 NC_016828 NC_00394 NC_004944 NC_004944 NC_004944 NC_004944 NC_004942 NC_014558 NC_010621 NC_014558 NC_016698 NC_017479 NC_004530 NC_004530 NC_017479 NC_017480 NC_016608 NC_016608 NC_017017	L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 1.00\\ 0.36\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.50\\ 1.00\\ 0.47\\ 1.00\\ 0.70\\ 0.67\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015591 NC.015597 NC.017324 NC.017327 NC.018683 NC.017327 NC.018683 NC.008717 NC.008524 NC.008739 I5 stability meat NC.014533 NC.010853 NC.014533 NC.014508 NC.014508 NC.015319 NC.008498 NC.008499 NC.012550 NC.015421	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquacolei sure : 0.754515 B. ccreus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis L. brevis L. brevis L. brevis L. brevis L. buevis	αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.00 0.23 0.23 0.23 0.25	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.004944 NC.004944 NC.014558 NC.010621 NC.014558 NC.016638 NC.015290 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.017479 NC.017480 NC.017480 NC.017480 NC.016608 NC.016608 NC.016636 NC.017017 NC.017018 NC.01277	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.50\\ 0.47\\ 1.00\\ 0.67\\ 0.50\\ 0.47\\ 1.00\\ 0.67\\ 1.00\\ 0.47\\ 1.00\\ 0.00\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013845 NC.015591 NC.017327 NC.017327 NC.018682 NC.017327 NC.018683 NC.009717 NC.006824 NC.008739 I5 stability meat NC.014333 NC.010853 NC.010853 NC.013533 NC.015319 NC.008498 NC.012550 NC.015421 NC.015429	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti M. avotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis L. brevis L. brevis L. brevis L. buchneri L. buchneri	αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli <th>0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.00 0.23 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 nan 0.27 0.00 0.47</th> <th>PL PL PL PL PL PL PL PL PL PL PL PL PL P</th> <th>NC.017498 NC.017500 NC.017500 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.006278 NC.014558 NC.014558 NC.014558 NC.016698 NC.015698 NC.015698 NC.017479 NC.017479 NC.017480 NC.017479 NC.017480 NC.017479 NC.017480 NC.016636 NC.017017 NC.016608 NC.016636 NC.017017 NC.012031</th> <th>L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. salivarius</th> <th>Bacilli Bacilli</th> <th>$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.47\\ 0.50\\ 0.47\\ 1.00\\ 0.00\\ 0.70\\ 0.50\\ 0.47\\ 1.00\\ 0.50\\$</th>	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.00 0.23 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 nan 0.27 0.00 0.47	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.017500 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.006278 NC.014558 NC.014558 NC.014558 NC.016698 NC.015698 NC.015698 NC.017479 NC.017479 NC.017480 NC.017479 NC.017480 NC.017479 NC.017480 NC.016636 NC.017017 NC.016608 NC.016636 NC.017017 NC.012031	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.47\\ 0.50\\ 0.47\\ 1.00\\ 0.00\\ 0.70\\ 0.50\\ 0.47\\ 1.00\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.013565 NC.013545 NC.015597 NC.017324 NC.017327 NC.017327 NC.018682 NC.018682 NC.008760 NC.008760 NC.008760 NC.008739 Is stability mean NC.010853 NC.001533 NC.0105510 NC.013533 NC.015339 NC.008498 NC.008499 NC.008499 NC.015421 NC.015421 NC.015429 NC.018611	S. fredii S. medicae S. medicae S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphtalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis L. brevis L. brevis L. buchneri L. buchneri L. buchneri L. buchneri	αproteobacteria βproteobacteria βproteobacteria βproteobacteria γproteobacteria βproteobacteria βproteobacteria γproteobacteria βproteobacteria βeailli Bacilli Ba	0.87 0.67 nan 1.00 0.33 1.00 0.36 0.33 1.00 0.33 1.00 0.23 0.23 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.36 0.00 0.36 0.00 0.36 0.00 0.36 0.00 0.36 0.32 0.25	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.006278 NC.006278 NC.01621 NC.015698 NC.01661 NC.015698 NC.016529 NC.004942 NC.004942 NC.004942 NC.004942 NC.004530 NC.017480 NC.017480 NC.017480 NC.017618 NC.016636 NC.016636 NC.016636 NC.017017 NC.012031 NC.012373	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius S. saureus	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.43\\ 1.00\\ 0.36\\ 0.47\\ 0.50\\ 1.00\\ 0.00\\ 0.47\\ 1.00\\ 0.67\\ 1.00\\ 0.50\\ 1.00\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.009622 NC.010865 NC.013545 NC.0135597 NC.017324 NC.017327 NC.018682 NC.017327 NC.018683 NC.009717 NC.008760 NC.008760 NC.008760 NC.008760 NC.008760 NC.008760 NC.008739 I5 stability mean NC.014333 NC.010853 NC.010853 NC.0105010 NC.013533 NC.014508 NC.013539 NC.008499 NC.008499 NC.015421 NC.015421 NC.015421 NC.018611 NC.018618	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis L. brevis L. brevis L. brevis L. buchneri L. buchneri L. carnosum	αproteobacteria βproteobacteria βproteobacteria βproteobacteria γproteobacteria βproteobacteria βacilli Bacilli	0.87 0.67 nan 1.00 0.33 1.00 0.33 1.00 0.33 0.33 0.23 0.23 0.25 0.73 1.00 0.33 0.32 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.36 0.00 0.36 0.00 0.37 1.00 0.36 0.33 0.32 0.25	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.006278 NC.006278 NC.010621 NC.015698 NC.015698 NC.015698 NC.015698 NC.016529 NC.004530 NC.004942 NC.004529 NC.004530 NC.017479 NC.017480 NC.017480 NC.017618 NC.016366 NC.017017 NC.016366 NC.017018 NC.01277 NC.01231 NC.013373 NC.016750	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. salivarius S. aureus S. macedonicus	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.50\\ 1.00\\ 0.70\\ 0.50\\ 1.00\\ 0.67\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.003037 NC.010865 NC.013545 NC.015591 NC.015597 NC.017324 NC.017327 NC.018682 NC.018683 NC.008717 NC.008683 NC.008760 NC.008760 NC.005241 NC.008760 NC.005241 NC.008739 I5 stability mear NC.014333 NC.010853 NC.010853 NC.013533 NC.014508 NC.015319 NC.015319 NC.015450 NC.015421 NC.015429 NC.015429 NC.018698 NC.018698 NC.018698 NC.018698	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis L. brevis L. brevis L. brevis L. brevis L. buchneri L. buchneri L. carnosum L. carnosum	αproteobacteria βproteobacteria βacilli Bacilli Bacilli<	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.23 0.23 0.23 0.25 0.25 0.73 1.00 0.36 0.00 nan 0.27 0.00 nan 0.27 0.00 nan 0.27 nan 0.47 0.47 0.47 0.00 0.57	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.004944 NC.004944 NC.004942 NC.014558 NC.016698 NC.015701 NC.013200 NC.004529 NC.006530 NC.017479 NC.017480 NC.017479 NC.017480 NC.016636 NC.016636 NC.016636 NC.017017 NC.012031 NC.01277 NC.012031 NC.016750 NC.016525	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. reuteri L. sakioarius L. salivarius L. salivarius S. macedonicus T. halophilus	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.33\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.50\\ 1.00\\ 0.70\\ 0.50\\ 1.00\\ 0.70\\ 0.50\\ 1.00\\ 0.67\\ 0.50\\ 1.00\\ 0.50\\ 0.50\\ 1.00\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.009622 NC.013565 NC.013545 NC.015597 NC.015597 NC.017324 NC.017324 NC.017327 NC.018682 NC.018683 NC.009717 NC.006824 NC.008241 NC.008241 NC.008739 I5 stability mear NC.014333 NC.010853 NC.010853 NC.013533 NC.014508 NC.008498 NC.008498 NC.0012550 NC.015421 NC.018698 NC.018699 NC.011352	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis L. brevis L. brevis L. brevis L. brevis L. buchneri L. buchneri L. carnosum L. carnosum L. carnosum	αproteobacteria βproteobacteria βacilli Bacilli Bacilli<	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.23 0.23 0.23 0.25 0.25 0.73 1.00 0.36 0.25 0.25	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.004944 NC.004944 NC.014558 NC.010621 NC.014558 NC.016619 NC.015200 NC.004942 NC.006529 NC.004942 NC.017479 NC.017480 NC.017479 NC.017480 NC.017480 NC.016636 NC.016636 NC.016636 NC.016636 NC.017018 NC.012031 NC.012031 NC.01233 NC.016750 NC.016233 NC.016253	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. salivarius L. salivarius S. sacedonicus T. halophilus T. halophilus	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.50\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.50\\ 1.00\\ 0.67\\ 1.00\\ 0.67\\ 1.00\\ 0.67\\ 1.00\\ 0.67\\ 1.00\\ 0.67\\ 1.00\\ 0.50\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.016814 NC.009621 NC.009622 NC.009622 NC.010865 NC.013845 NC.013545 NC.017327 NC.017327 NC.017327 NC.018682 NC.009717 NC.006824 NC.005241 NC.008739 E5 stability meat NC.014333 NC.010853 NC.010853 NC.0105510 NC.015319 NC.008498 NC.012550 NC.015421 NC.018699 NC.018699 NC.011352 NC.014592	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. brevis L. brevis L. brevis L. brevis L. buchneri L. buchneri L. carnosum L. casei L. casei L. casei L. casei	$\begin{array}{c} \alpha \text{proteobacteria} \\ \beta proteobacte$	0.87 0.67 nan 1.00 0.00 0.33 1.00 0.36 0.33 1.00 0.02 0.23 0.23 0.25 0.73 1.00 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 0.36 0.00 0.36 0.27 0.43 1.00 0.27 0.27 0.26	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.017500 NC.016820 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.004944 NC.004944 NC.014558 NC.010621 NC.014558 NC.016618 NC.015200 NC.004942 NC.006529 NC.017479 NC.017479 NC.017480 NC.017479 NC.017480 NC.017479 NC.017480 NC.017479 NC.016636 NC.017017 NC.016636 NC.017017 NC.016636 NC.016636 NC.017017 NC.012031 NC.013273 NC.016520 NC.015225 NC.015255 NC.015255 NC.015260	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. salivarius	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.30\\ 0.47\\ 1.00\\ 0.30\\ 0.47\\ 1.00\\ 0.00\\ 0.00\\ 0.47\\ 1.00\\ 0.50\\ 1.00\\ 0.47\\ 1.00\\ 0.50\\ 0.47\\ 1.00\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL	NC.016814 NC.009621 NC.009622 NC.003037 NC.013565 NC.013545 NC.013591 NC.015597 NC.017324 NC.017327 NC.018682 NC.018682 NC.008760 NC.008717 NC.006824 NC.008760 NC.005241 NC.008760 NC.005241 NC.008739 E5 stability meas NC.014533 NC.010853 NC.010533 NC.015333 NC.015319 NC.008498 NC.015429 NC.015429 NC.018698 NC.018698 NC.018698 NC.014528 NC.014528 NC.014528 NC.014529 NC.018698 NC.01352 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014528 NC.014619 NC.004528 NC.01467	S. fredii S. medicae S. medicae S. meliloti S. meliloti M. avotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. brevis L. brevis L. brevis L. brevis L. buchneri L. buchneri L. buchneri L. curnosum L. carnosum L. casei L. citreum	αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli <th>0.87 0.67 nan 1.00 0.33 1.00 0.33 1.00 0.33 1.00 0.33 1.00 0.23 0.23 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 nan 0.27 0.00 0.43 1.00 0.36 1.00 0.27 0.36 1.00</th> <th>PL PL PL PL PL PL PL PL PL PL PL PL PL P</th> <th>NC.017498 NC.017500 NC.017500 NC.016820 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.006278 NC.014558 NC.016621 NC.014558 NC.016621 NC.014558 NC.016639 NC.017479 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017479 NC.017480 NC.017017 NC.012031 NC.012031 NC.010523 NC.015255 NC.015257 NC.015257 NC.015261</th> <th>L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. salivariu</th> <th>Bacilli Bacilli</th> <th>$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.67\\ 0.50\\ 1.00\\ 0.67\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\$</th>	0.87 0.67 nan 1.00 0.33 1.00 0.33 1.00 0.33 1.00 0.33 1.00 0.23 0.23 0.23 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.00 nan 0.27 0.00 0.43 1.00 0.36 1.00 0.27 0.36 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.017500 NC.016820 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.004944 NC.006278 NC.014558 NC.016621 NC.014558 NC.016621 NC.014558 NC.016639 NC.017479 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017480 NC.017479 NC.017480 NC.017017 NC.012031 NC.012031 NC.010523 NC.015255 NC.015257 NC.015257 NC.015261	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. salivariu	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.43\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.80\\ 0.47\\ 1.00\\ 0.67\\ 0.50\\ 1.00\\ 0.67\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\ 1.00\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL	NC.016814 NC.009621 NC.009622 NC.009622 NC.013565 NC.013545 NC.0135597 NC.017324 NC.017327 NC.017327 NC.018682 NC.017327 NC.018682 NC.008760 NC.008760 NC.008760 NC.008760 NC.008739 Estability mean NC.008739 Stability mean NC.014533 NC.005010 NC.013533 NC.005010 NC.013533 NC.005010 NC.013533 NC.005010 NC.013533 NC.005010 NC.013533 NC.008499 NC.008499 NC.008499 NC.0015421 NC.015429 NC.018611 NC.018698 NC.018699 NC.01352 NC.014619 NC.004528 NC.010469	S. fredii S. medicae S. medicae S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti X. autotrophicus A. aromaticum P. naphthalenivorans R. eutropha M. aquaeolei sure : 0.754515 B. cereus B. coagulans E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. faecalis E. brevis L. brevis L. brevis L. buchneri L. buchneri L. carnosum L. carsei L. casei L. citreum	αproteobacteria βproteobacteria βproteobacteria βproteobacteria γproteobacteria βproteobacteria βproteobacteria γproteobacteria βproteobacteria βacilli Bacilli Ba	0.87 0.67 nan 1.00 0.33 1.00 0.36 0.33 1.00 0.33 1.00 0.23 0.23 0.23 0.23 0.25 0.25 0.73 1.00 0.36 0.00 0.36 0.00 0.36 0.00 0.57 nan 0.47 0.43 1.00 0.27 0.36 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.017498 NC.017500 NC.008496 NC.016820 NC.016821 NC.016828 NC.003894 NC.004944 NC.006278 NC.006278 NC.01621 NC.014558 NC.016621 NC.015698 NC.016619 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.004942 NC.017480 NC.017480 NC.017480 NC.017618 NC.016036 NC.016036 NC.016036 NC.016037 NC.016037 NC.012031 NC.013373 NC.016750 NC.015255 NC.015257 NC.015261	L. lactis L. lactis L. lactis L. mesenteroides L. mesenteroides L. mesenteroides L. mesenteroides L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. plantarum L. reuteri L. reuteri L. reuteri L. reuteri L. sakivarius L. salivarius L. sali	Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.73\\ 1.00\\ 0.36\\ 0.50\\ 1.00\\ 0.33\\ 0.43\\ 1.00\\ 0.33\\ 0.43\\ 1.00\\ 0.33\\ 0.47\\ 0.50\\ 1.00\\ 0.67\\ 1.00\\ 0.67\\ 1.00\\ 0.67\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 1.00\\ 0.50\\ 1.00\\$

PL	NC_015256	T. muriaticus	Bacilli	0.50					
PL	$NC_{-}015756$	W. koreensis	Bacilli	1.00	cluster	: 19 stability	measure : 0.832028		
PL	NC_010404	A. baumannii	γ proteobacteria	1.00	CHR	NC_017098	S. africana	Spirochaetes	1.00
cluster	• 16 stability n	1easure : 0.983268			CHR	NC-015732 NC 015436	S. caldaria S. coccoides	Spirochaetes	1.00
CHR	NC_004552	C. abortus	Chlamydiia	1.00	CHR	NC_015152	S. globus	Spirochaetes	0.50
CHR	NC_003361	C. caviae	Chlamydiia	1.00	CHR	NC_016633	S. pleomorpha	Spirochaetes	0.00
CHR	NC_007899	C. felis	Chlamydiia	1.00	CHR	NC_014364	S. smaragdinae	Spirochaetes	0.64
CHR	NC_002620	C. muridarum	Chlamydiia	1.00	CHR	NC_015577	T. azotonutricium T. hronnahononao	Spirochaetes	0.12
CHR	NC_000922	C. preumoniae	Chlamydija	1.00	CHR	NC_015714	T. paraluiscuniculi	Spirochaetes	1.00
CHR	NC_002179	C. pneumoniae	Chlamydiia	1.00	CHR	NC_015578	T. primitia	Spirochaetes	1.00
CHR	NC_002491	C. pneumoniae	Chlamydiia	1.00	CHR	NC_015385	T. succinifaciens	Spirochaetes	1.00
CHR	NC_005043	C. pneumoniae	Chlamydiia	1.00	CHR	NC_000919	T. pallidum	Spirochaetes	1.00
CHR	NC_017285	C. pneumoniae	Chlamydiia	1.00	CHR	NC_016844	T. pallidum	Spirochaetes	1.00
CHR	NC 014796	Cana. Frotochamydia	Chlamydiia	1.00	CHR	NC 016842	T. pallidum	Spirochaetes	1.00
CHR	NC_015470	C. psittaci	Chlamydiia	1.00	CHR	NC_016843	T. pallidum	Spirochaetes	1.00
CHR	NC_017287	C. psittaci	Chlamydiia	1.00	CHR	NC_016848	T. pallidum	Spirochaetes	1.00
CHR	NC_017289	C. psittaci	Chlamydiia	1.00	CHR	$NC_{-}017268$	T. pallidum	Spirochaetes	1.00
CHR	NC_017290	C. psittaci	Chlamydiia	1.00	CHR	NC_018722	T. pallidum	Spirochaetes	1.00
CHR	NC 017291	C. psittaci	Chlamydiia	1.00 nan	cluster	· 20 stability	measure : 0 806525		
CHR	NC_018619	C. psittaci	Chlamydiia	1.00	PL	NC_002146	B. anthracis	Bacilli	1.00
CHR	NC_018620	C. psittaci	Chlamydiia	1.00	PL	NC_003981	B. anthracis	Bacilli	1.00
CHR	NC_018621	C. psittaci	Chlamydiia	1.00	\mathbf{PL}	NC_007323	B. anthracis	Bacilli	1.00
CHR	NC_018622	C. psittaci	Chlamydiia	1.00	PL	NC_012577	B. anthracis	Bacilli	1.00
CHR	NC_018623 NC_018624	C. psittaci	Chlamydiia	1.00	PL	NC_012655	B. anthracis	Bacilli	1.00
CHR	NC_018625	C. psittaci	Chlamydija	1.00	PL PL	NC 014332	B. anthracis B. cereus	Bacilli	1.00
CHR	NC_018626	C. psittaci	Chlamydiia	1.00	PL	NC_018499	B. cereus	Bacilli	0.86
CHR	$NC_{-}018627$	C. psittaci	Chlamydiia	1.00	\mathbf{PL}	NC_015148	B. subtilis	Bacilli	1.00
CHR	NC_019391	C. psittaci	Chlamydiia	1.00	\mathbf{PL}	NC_{006578}	B. thuringiensis	Bacilli	1.00
CHR	NC_000117	C. trachomatis	Chlamydiia	1.00	PL	NC_008598	B. thuringiensis	Bacilli	0.36
CHR	NC 010280	C. trachomatis C. trachomatis	Chlamydiia	1.00	PL PL	NC_010283	B. thuringiensis B. thuringiensis	Bacilli	1.00
CHR	NC_010287	C. trachomatis	Chlamydiia	1.00	PL	NC_018502	B. thuringiensis	Bacilli	0.57
CHR	NC_012686	C. trachomatis	Chlamydiia	1.00	\mathbf{PL}	NC_018688	B. thuringiensis	Bacilli	0.53
CHR	NC_012687	C. trachomatis	Chlamydiia	1.00	\mathbf{PL}	NC_018686	B. thuringiensis	Bacilli	1.00
CHR	NC_015744	C. trachomatis	Chlamydiia	1.00	PL	NC_018687	B. thuringiensis	Bacilli	1.00
CHR	NC 017429	C. trachomatis	Chlamydiia	1.00	PL PL	NC 010181	B. weinenstephanensis	Bacilli	1.00
CHR	NC_017430	C. trachomatis	Chlamydiia	1.00	PL	NC_015390	Carnobacterium sp.	Bacilli	0.81
CHR	NC_017431	C. trachomatis	Chlamydiia	1.00	\mathbf{PL}	NC_010608	E. arabatum	Bacilli	1.00
CHR	NC_017432	C. trachomatis	Chlamydiia	1.00	PL	NC_002630	E. faecalis	Bacilli	0.29
CHR	NC_017434	C. trachomatis	Chlamydiia	1.00	PL	NC_004669	E. faecalis	Bacilli	0.50
CHR	NC_017430	C. trachomatis	Chlamydiia	1.00	PL PL	NC 006827	E. jaecalis E. faecalis	Bacilli	0.59
CHR	NC_017439	C. trachomatis	Chlamydiia	1.00	PL	NC_008445	E. faecalis	Bacilli	1.00
CHR	NC_017440	C. trachomatis	Chlamydiia	1.00	\mathbf{PL}	$NC_{-}013514$	E. faecalis	Bacilli	1.00
CHR	NC_017441	C. trachomatis	Chlamydiia	1.00	PL	NC_014475	E. faecalis	Bacilli	1.00
CHR	NC 017951 NC 017952	C. trachomatis C. trachomatis	Chlamydiia	1.00	PL PL	NC_014726 NC_018222	E. faecalis E. faecalis	Bacilli	1.00
CHR	NC_017953	C. trachomatis	Chlamydiia	nan	PL	NC_018223	E. faecalis	Bacilli	1.00
CHR	$NC_{-}015702$	$P.\ a can that oe bae$	Chlamydiia	1.00	\mathbf{PL}	$NC_{-}007594$	E. faecium	Bacilli	1.00
CHR	NC_015713	S. negevensis	Chlamydiia	1.00	PL	NC_008768	E. faecium	Bacilli	1.00
DI	NC_014225	W. chondrophila	Chlamydiia	1.00	PL	NC_008821	E. faecium	Bacilli	1.00
FL	NC_014749	C. nitroreaucens	Deferribacteres	0.27	PL PL	NC 016967	E. faecium E. faecium	Bacilli	1.00
cluster	: 17 stability n	neasure : 0.778260			PL	NC_017961	E. faecium	Bacilli	0.94
CHR	NC_011146	G. bemidjiensis	δ proteobacteria	1.00	\mathbf{PL}	NC_{017963}	E. faecium	Bacilli	1.00
CHR	NC_011979	G. daltonii	δ proteobacteria	1.00	PL	NC_010880	E. faecium	Bacilli	0.95
DI	NC_010814	G. lovleyi	∂proteobacteria	1.00	PL	NC_010980	E. faecium	Bacilli	1.00
CHR	NC_007517	G. metallireducens	δproteobacteria	1.00	PL	NC_011364	E. faecium	Bacilli	1.00
CHR	NC_002939	G. sulfurreducens	δ proteobacteria	1.00	\mathbf{PL}	NC_017032	E. faecium	Bacilli	1.00
CHR	NC_017454	$G. \ sulfur educens$	δ proteobacteria	1.00	\mathbf{PL}	NC_015845	E. hirae	Bacilli	1.00
CHR	NC_009483	G. uraniireducens	δ proteobacteria	1.00	PL	NC_012551	L. brevis	Bacilli	1.00
CHR	NC 014973	Geobacter sp.	oproteobacteria	1.00	PL PL	NC_015420 NC_008502	L. buchneri L. casei	Bacilli	1.00
CHR	NC_007498	P. carbinolicus	δproteobacteria	1.00	PL	NC_017475	L. casei	Bacilli	0.45
PL	NC_008607	P. propionicus	δ proteobacteria	0.12	PL	NC_017476	L. casei	Bacilli	0.55
PL	NC_008608	P. propionicus	δ proteobacteria	1.00	\mathbf{PL}	$NC_{-}018674$	L. carnosum	Bacilli	0.50
CHR	NC_008609	P. propionicus	δ proteobacteria	1.00	PL	NC_018675	L. carnosum	Bacilli	0.55
CHR	NC_007759	S. aciditrophicus	oproteobacteria	0.59	PL	NC_010466	L. citreum	Bacilli	0.36
cluster	: 18 stability n	neasure : 0.780123			PL	NC_010540	L. garvieae	Bacilli	1.00
CHR	NC_014844	D. aespoeensis	δ proteobacteria	1.00	\mathbf{PL}	NC_014496	L. grayi	Bacilli	1.00
CHR	NC_016629	D. africanus	δ proteobacteria	1.00	\mathbf{PL}	NC_003383	L. innocua	Bacilli	1.00
CHR	NC_007519	D. alaskensis	∂proteobacteria	1.00	PL	NC_016827	L. mesenteroides	Bacilli Ba-:11:	0.67
CHR	NC 011892	D. desulfuricans	oproteobacteria	1.00	PL PI	NC 014255	ь. monocytogenes	Bacilli	0.67
CHR	NC_016803	D. desulfuricans	δproteobacteria	1.00	PL	NC_014495	L. monocytogenes	Bacilli	0.50
CHR	NC_012796	D. magneticus	δ proteobacteria	1.00	PL	NC_018888	L. monocytogenes	Bacilli	1.00
CHR	NC_013223	D. retbaense	δ proteobacteria	1.00	\mathbf{PL}	NC_018889	$L.\ monocytogenes$	Bacilli	1.00
CHR	NC_002937	D. vulgaris	δ proteobacteria	1.00	PL	NC_011225	L. rhamnosus	Bacilli	1.00
CHR	NC_008751	D. vulgaris	oproteobacteria	1.00	PL DT	NC_015980	L. santranciscensis	Bacilli	0.36
CHR	NC_011769 NC_017310	D. vulgaris	oproteobacteria δproteobacteria	1.00	PL PL	NC_011995 NC_011996	M. caseolyticus M. caseoluticus	Bacilli	1.00
CHR	NC_012881	D. salexigens	δ proteobacteria	1.00	PL	NC_015517	M. plutonius	Bacilli	0.67
CHR	NC_008011	$L.\ intracellularis$	δ proteobacteria	1.00	\mathbf{PL}	NC_018265	M. plutonius	Bacilli	1.00
PL	NC_008012	L. intracellularis	δ proteobacteria	0.62	PL	NC_010715	N. thermophilus	Clostridia	0.14
PL	NC_003241	Nostoc sp.	нотодопае	1.00	PL	110_010864	r. aciailactici	Baciili	1.00

Annexe D.Résultats du clustering de V^R par INFOMAP

PL	NC 016607	P claussenii	Bacilli	1 00	CHB	NC 017847	A haumannii	vproteobacteria	1.00
PL	NC_017019	P. claussenii	Bacilli	0.96	CHR	NC_018706	A. baumannii	γ proteobacteria	1.00
PL	NC_002136	S. agalactiae	Bacilli	1.00	CHR	NC_016603	A. calcoaceticus	γ proteobacteria	1.00
PL	NC_013453	S. aureus	Bacilli	0.93	CHR	NC_014259	A. oleivorans	γ proteobacteria	1.00
PL	NC_005173	S. pyogenes S. simulans	Bacilli	1.00	Спк	NC-005966	Acinetobacier sp.	γproteobacteria	1.00
\mathbf{PL}	NC_013944	S. simulans	Bacilli	1.00	cluster :	23 stability m	neasure : 0.804747		
PL	NC_012923	S. suis	Bacilli	1.00	\mathbf{PL}	NC_017908	M. abscessus	Actinobacteridae	1.00
clustor	· 21 stability	$m_{0.05} = 0.080027$			PL	NC_016030	G. xylinus	αproteobacteria	0.43
PL	NC_013548	R. palustris	αproteobacteria	1.00	ГL	01000123	N. pentaromatioorans	aproteobacteria	0.13
PL	NC_009793	C. koseri	γ proteobacteria	1.00	\mathbf{PL}	NC_009432	R. sphaeroides	α proteobacteria	0.38
PL	NC_003114	C. rodentium	γ proteobacteria	1.00	RECE	NC_014013	S. japonicum	α proteobacteria	0.18
PL	NC_004940	E. amylovora E. anli	γ proteobacteria	1.00	PL	NC_007353	Sphingomonas sp.	α proteobacteria	1.00
PL	NC 019079	E. coli E. coli	γproteobacteria	1.00	PL	NC 014911	A. denitrificans	βproteobacteria	1.00
PL	NC_019077	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_006830	A. xylosoxidans	β proteobacteria	1.00
PL	NC_019076	E. coli	γ proteobacteria	1.00	PL	NC_014641	A. xylosoxidans	β proteobacteria	1.00
PL	NC_019056	E. coli E. coli	γ proteobacteria	1.00	PL	NC-008766	Acidovorax sp.	β proteobacteria	1.00
PL	NC_018997	E. coli E. coli	γ proteobacteria γ proteobacteria	1.00	PL	NC_008385	B. cepacia	βproteobacteria	1.00
\mathbf{PL}	NC_017721	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_008459	B. pertussis	β proteobacteria	1.00
\mathbf{PL}	NC_{017662}	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_013191	Cand. Accumulibacter	β proteobacteria	0.57
PL	NC_017661	E. coli	γ proteobacteria	1.00	PL	NC_013193	Cand. Accumulibacter	β proteobacteria	1.00
PL PL	NC 017654	E. coli E. coli	vproteobacteria	1.00	PL	NC 016968	C. testosteroni C. testosteroni	βproteobacteria	1.00
PL	NC_017636	E. coli	γ proteobacteria	1.00	PL	NC_016978	$C. \ testos teroni$	β proteobacteria	1.00
PL	NC_016903	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_005088	D. acidovorans	β proteobacteria	1.00
PL	NC_014543	E. coli	γ proteobacteria	1.00	PL	NC_012970	M. glucosetrophus	β proteobacteria	1.00
PL PL	NC 013368	E. coli E. coli	γproteobacteria	1.00	PL PL	NC 017510	N. gonorrhoeae N. gonorrhoeae	βproteobacteria	1.00 nan
PL	NC_013363	E. coli	γ proteobacteria	1.00	PL	NC_005912	R. eutropha	β proteobacteria	1.00
PL	NC_011799	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_007337	$R. \ eutropha$	β proteobacteria	1.00
PL	NC_011408	E. coli	γ proteobacteria	1.00	PL	NC_010606	A. baumannii	γ proteobacteria	0.80
PL PL	NC 010485	E. coli E. coli	γproteobacteria	1.00	PL PL	NC 017166	A. baumannii A baumannii	γ proteobacteria	1.00
PL	NC_008488	E. coli	γ proteobacteria	1.00	PL	NC_017848	A. baumannii	γ proteobacteria	0.80
PL	$NC_{-}005970$	E. coli	γ proteobacteria	1.00	\mathbf{PL}	NC_006143	A. caviae	γ proteobacteria	1.00
PL	NC_005019	E. coli	γ proteobacteria	1.00	PL	NC_010919	A. hydrophila	γ proteobacteria	1.00
PL PL	NC 001371	E. coli E. coli	γ proteobacteria γ proteobacteria	1.00	PL PL	NC 013509	E. aerogenes E. tarda	γ proteobacteria	1.00
PL	NC_002497	E. ictaluri	γ proteobacteria	1.00	PL	NC_017508	M. adhaerens	γ proteobacteria	1.00
\mathbf{PL}	NC_017391	E. pyrifoliae	γ proteobacteria	1.00	\mathbf{PL}	NC_017858	Methylophaga sp.	γ proteobacteria	1.00
PL	NC_017388	E. pyrifoliae	γ proteobacteria	1.00	PL	NC_008357	P. aeruginosa	γ proteobacteria	1.00
PL PL	NC_013954 NC 013264	E. pyrifoliae E. pyrifoliae	γ proteobacteria γ proteobacteria	1.00	PL PL	NC 012919 NC 013176	P. damselae P. nutida	γproteobacteria	1.00
PL	NC_015515	E. sp.	γ proteobacteria	1.00	PL	NC_004956	Pseudomonas sp.	γ proteobacteria	1.00
PL	NC_009716	E. sp.	γ proteobacteria	0.88	\mathbf{PL}	NC_017555	X. albilineans	γ proteobacteria	1.00
PL	NC_004936	E. sp.	γ proteobacteria	1.00	PL	NC_017556	X. albilineans	γ proteobacteria	0.71
PL PL	NC_019159 NC_019156	K. pneumoniae K. pneumoniae	γ proteobacteria γ proteobacteria	1.00	PL PL	NC 010579	X. fastidiosa X. fastidiosa	γproteobacteria	1.00
PL	NC_018953	K. pneumoniae	γproteobacteria	nan	PL	NC_003430	uncultured	Environment	1.00
PL	NC_{016847}	K. pneumoniae	γ proteobacteria	1.00	\mathbf{PL}	NC_004840	uncultured	Environment	1.00
PL	NC_011640	K. pneumoniae	γ proteobacteria	1.00	PL	NC_006352	uncultured	Environment	1.00
PL PL	NC_011382 NC_009653	K. pneumoniae K. pneumoniae	γ proteobacteria	1.00	PL PL	NC_007680	uncultured	Environment	1.00
PL	NC_009652	K. pneumoniae	γproteobacteria	1.00	PL	NC_019021	uncultured	Environment	1.00
PL	$NC_{-}005015$	K. pneumoniae	γ proteobacteria	1.00	\mathbf{PL}	NC_019022	uncultured	Environment	1.00
PL	NC_002610	K. pneumoniae	γ proteobacteria	1.00		04	0.010500		
PL PL	NC 002632	P. atrosepticum P. vulgaris	γ proteobacteria γ proteobacteria	1.00	CHB	24 stability m	M gallisenticum	Mollicutes	1.00
PL	NC_019136	S. enterica	γ proteobacteria	1.00	CHR	NC_017502	M. gallisepticum	Mollicutes	1.00
PL	NC_019102	S. enterica	γ proteobacteria	1.00	CHR	NC_017503	M. gallisepticum	Mollicutes	1.00
PL	NC_015575	S. enterica	γ proteobacteria	1.00	CHR	NC_018406	M. gallisepticum	Mollicutes	nan
PL	NC_015570	S. enterica	γproteobacteria	1.00	CHR	NC_018407 NC_018408	M. gaussepticum	Mollicutes	1.00
PL	NC_011214	S. enterica	γ proteobacteria	1.00	CHR	NC_018409	M. gallisepticum	Mollicutes	1.00
PL	NC_011082	S. enterica	γ proteobacteria	1.00	CHR	NC_018413	M. gallisepticum	Mollicutes	1.00
PL PL	NC-006815 NC 003457	S. enterica S. enterica	γ proteobacteria γ proteobacteria	1.00	CHR	NC_018410	M. gallisepticum	Mollicutes Mollicutes	1.00
PL	NC_003079	S. enterica	γproteobacteria	1.00	CHR	NC_018412	M. gallisepticum	Mollicutes	1.00
\mathbf{PL}	NC_017330	S. flexneri	γ proteobacteria	1.00	CHR	NC_000908	M. genitalium	Mollicutes	1.00
PL	NC_002773	S. flexneri	γ proteobacteria	1.00	CHR	NC_018497	M. genitalium	Mollicutes	1.00
PL PL	NC_016823 NC_009346	S. sonnei S. sonnei	γ proteobacteria	1.00	CHR CHB	NC_018495	M. genitalium M. genitalium	Mollicutes	0.64
PL	NC_017264	Y. pestis	γ proteobacteria	1.00	CHR	NC_018498	M. genitalium	Mollicutes	0.50
PL	NC_015055	Y. pestis	γ proteobacteria	1.00	CHR	NC_000912	M. pneumoniae	Mollicutes	1.00
PL	NC_008121	Y. pestis	γ proteobacteria	1.00	CHR	NC_016807	M. pneumoniae	Mollicutes	1.00
PL PL	NC_008119 NC_005816	1. pestis Y. pestis	γproteobacteria γproteobacteria	1.00	CHR CHB	NC_017504	M. pneumoniae M. penetrans	Mollicutes	1.00
PL	NC_004837	Y. pestis	γ proteobacteria	1.00	Cint	1.0.004432	penetrano	monicutes	1 0.13
PL	$NC_{-}003132$	Y. pestis	γ proteobacteria	nan	cluster	25 stability m	neasure : 0.795510		
-l'					PL	NC_001496	B. anthracis	Bacilli	1.00
CHR	: 22 stability	A haumannii	oproteobactoria	1.00	PL DI	NC_003980	B. anthracis B. anthracis	Bacilli Bacilli	1.00
CHR	NC_010400	A. baumannii	γproteobacteria	1.00	PL	NC_012656	B. anthracis	Bacilli	1.00
CHR	NC_010410	A. baumannii	γ proteobacteria	1.00	PL	NC_012579	B. anthracis	Bacilli	1.00
CHR	NC_010611	A. baumannii	γ proteobacteria	1.00	PL	NC_017726	B. anthracis	Bacilli	1.00
CHR	NC_011586 NC 011595	A. baumannıı A. baumannii	γ proteobacteria γ proteobacteria	1.00	PL DI	NC_005707	B. cereus	Bacilli	1.00
CHR	NC_017162	A. baumannii	γ proteobacteria	1.00	PL	NC_007105	B. cereus	Bacilli	0.36
CHR	NC_017171	A. baumannii	γ proteobacteria	1.00	\mathbf{PL}	NC_010916	B. cereus	Bacilli	1.00
CHR	NC_017387	A. baumannii	γ proteobacteria	nan	PL	NC_010921	B. cereus	Bacilli	1.00
					PL	INC_010924	D. cereus	Daciili	1.00

		-		1			-		1
PL	NC_010933	B. cereus	Bacilli	0.94	PL	NC_015727	C. necator	β proteobacteria	1.00
PL	NC_010934	B. cereus	Bacilli	1.00	\mathbf{PL}	NC_010529	$C. \ taiwanensis$	β proteobacteria	1.00
PL	NC_011337	B. cereus	Bacilli	0.94	\mathbf{PL}	NC_008757	P. naphthalenivorans	β proteobacteria	1.00
PL	NC_011342	B. cereus	Bacilli	0.00	\mathbf{PL}	NC_008759	P. naphthalenivorans	β proteobacteria	1.00
PL	NC_011339	B. cereus	Bacilli	1.00	\mathbf{PL}	NC_007949	Polaromonas sp.	β proteobacteria	0.73
PL	NC_011656	B. cereus	Bacilli	1.00	\mathbf{PL}	NC_007950	Polaromonas sp.	β proteobacteria	1.00
PL	NC_011655	B. cereus	Bacilli	1.00	\mathbf{PL}	NC_007901	R. ferrireducens	β proteobacteria	1.00
PL	NC_011775	B. cereus	Bacilli	0.32	\mathbf{PL}	NC_012849	R. pickettii	β proteobacteria	1.00
PL	NC_011777	B. cereus	Bacilli	1.00	\mathbf{PL}	NC_012855	R. pickettii	β proteobacteria	1.00
LP	NC_011973	B. cereus	Bacilli	1.00	\mathbf{PL}	NC_019151	K. pneumoniae	γ proteobacteria	1.00
PL	NC 012473	B cereus	Bacilli	1.00		1	F	/	1
PL	NC 014331	B cereus	Bacilli	1.00	cluster	· 28 stability n	0.2850 0.742857		
DI	NC 014757	P acress	Pagilli	1.00	CUD	NC 019492	A	Asidahastania	1.00
	NC-014757	B. cereus	Dacilli	1.00	CHR	NC_012483	A. capsulatum	Acidobacteria	1.00
	NC_016792	B. cereus	Bacilli	1.00	CHR	NC_008536	C. Solibacter	Solibacteres	0.04
PL	NC_018492	B. cereus	Bacilli	1.00	CHR	NC_008009	Cand. Koribacter	Unidentified	1.00
PL	NC_018493	B. cereus	Bacilli	1.00	CHR	NC_016631	G. mallensis	Acidobacteria	1.00
PL	NC_004604	B. megaterium	Bacilli	0.00	CHR	NC_015064	G. tundricola	Acidobacteria	1.00
PL	NC_014023	B. megaterium	Bacilli	0.00	\mathbf{PL}	NC_015065	G. tundricola	Acidobacteria	1.00
PL	NC_014025	B. megaterium	Bacilli	0.00	\mathbf{PL}	NC_015057	G. tundricola	Acidobacteria	0.43
PL	NC_017139	B. megaterium	Bacilli	0.97	CHB	NC 014963	T saanensis	Acidobacteria	1 00
PL	NC 014937	B thuringiensis	Bacilli	1.00	CHB	NC 018014	T roseus	Acidobacteria	1.00
PI.	NC 009841	B thuringiensis	Bacilli	0.63	01110	1 1101010011	11.100000	Therefore a second	1 1.00
PL	NC 018870	B. thuringiansis	Bacilli	0.50	aluston	. 20 stability w	0.068056		
	NC 018501	D. thuringtensis	Dacilli	1.00	Cluster	29 stability in	1easure : 0.908050		1.00
	NC_014179	B. inuringiensis	Dacilli	1.00	CHR	NC_014761	O. profundus	Deinococci	1.00
PL	NC_014172	B. thuringiensis	Bacilli	1.00	CHR	NC_015387	M. hydrothermalis	Deinococci	1.00
PL	NC_017199	B. thuringiensis	Bacilli	0.70	CHR	NC_013946	M. ruber	Deinococci	1.00
PL	NC_018880	B. thuringiensis	Bacilli	0.19	CHR	NC_014212	M. silvanus	Deinococci	1.00
PL	NC_010076	B. thuringiensis	Bacilli	0.95	CHR	NC_019386	T. oshimai	Deinococci	1.00
PL	NC_018685	B. thuringiensis	Bacilli	0.94	CHR	NC_014221	T. radiovictrix	Deinococci	0.43
PL	NC_018488	B. thuringiensis	Bacilli	0.62	CHR	NC_014974	T. scotoductus	Deinococci	1.00
PL	NC_018489	B. thuringiensis	Bacilli	0.46	CHB	NC 005835	T thermophilus	Deinococci	1.00
PL	NC 018486	B. thuringiensis	Bacilli	1.00	СНБ	NC 006461	T thermonhilus	Deinococci	1.00
PI.	NC 017202	B thuringianeis	Bacilli	1.00	CUD	NC 017070	T theme Lil.	Deinococci	1.00
DT	NC 017202	B. thuringiensis	Dacilli	1.00	CHK	NG_017272	1. inermophilus	Demococci	1.00
	NC_017202	B. inuringiensis	Bacilli	nan	CHR	NC_017587	T. thermophilus	Deinococci	1.00
	NC_017201	B. thuringiensis	Bacilli	0.36	CHR	NC_017278	Thermus sp.	Deinococci	1.00
PL	NC_017206	B. thuringiensis	Bacilli	0.43					
PL	NC_017205	B. thuringiensis	Bacilli	0.75	cluster	: 30 stability m	1easure : 0.838174		
PL	NC_017212	B. thuringiensis	Bacilli	0.62	RECE	NC_006933	B. abortus	α proteobacteria	1.00
PL	NC_013963	Bacillus sp.	Bacilli	1.00	RECE	NC_010740	B. abortus	α proteobacteria	1.00
PL	NC_015661	G. thermoglucosidasius	Bacilli	0.37	RECE	NC_016777	B. abortus	α proteobacteria	1.00
PL	NC_013317	S. aureus	Bacilli	0.83	RECE	NC_010104	B. canis	α proteobacteria	1.00
					RECE	NC 016796	B. canis	αproteobacteria	1.00
cluster :	26 stability me	asure : 0.942302			BECE	NC 003318	B melitensis	oproteobacteria	1.00
RECE	NC 006841	A. fischeri	γ proteobacteria	1.00	BECE	NC 007624	B melitensis	aproteobacteria	1.00
BECE	NC 011186	A fischeri	oproteobacteria	1.00	PECE	NC 012442	P. melitensis	aproteobacteria	1.00
RECE	NC 011313	A salmonicida	oproteobacteria	1.00	DECE	NC 017945	D. memensis	aproteobacteria	1.00
DECE	NC 015627	V an avillanum	proteobacteria	1.00	RECE	NC_017245	B. metitensis	aproteobacteria	1.00
NECE	NC_015057	v. angumaram	γproteobacteria	1.00	RECE	NC_017247	B. melitensis	αproteobacteria	1.00
RECE	NC_012667	V. cholerae	γ proteobacteria	1.00	RECE	NC_017283	B. melitensis	α proteobacteria	1.00
RECE	NC_016945	V. cholerae	γ proteobacteria	1.00	RECE	NC_013118	B. microti	α proteobacteria	1.00
RECE	NC_009456	V. cholerae	γ proteobacteria	1.00	RECE	NC_009504	B. ovis	α proteobacteria	1.00
RECE	NC_012580	V. cholerae	γ proteobacteria	1.00	RECE	NC_015858	B. pinnipedialis	α proteobacteria	1.00
RECE	NC_017269	V. cholerae	γ proteobacteria	1.00	RECE	NC_004311	B. suis	α proteobacteria	1.00
RECE	NC_012583	V. cholerae	γ proteobacteria	1.00	RECE	NC_010167	B. suis	α proteobacteria	1.00
RECE	NC_002506	V. cholerae	γ proteobacteria	1.00	BECE	NC 016775	B suis	oproteobacteria	1.00
RECE	NC_016446	V. cholerae	γ proteobacteria	1.00	BECE	NC 017250	B suis	oproteobacteria	1.00
BECE	NC 016628	V furnissii	oproteobacteria	1.00	BECE	NC 009668	O anthroni	oproteobactoria	0.05
RECE	NC 009784	V harveni	oproteobacteria	nan	DI	NC 007506	V. aampastria	aproteobacteria	0.35
PECE	NC 004605	V. narobacmolations	(proteobacteria	1.00	ГL	1001007500	A. campestris	proteobacteria	0.00
DECE	NC 011744	V. paranaemoigneas	yproteobacteria	1.00		0.1	0.000151		
RECE	NC_011744	V. spienaiaus	γ proteobacteria	1.00	cluster	: 31 stability m	1easure : 0.869474		
RECE	NC_004460	V. vulnificus	γ proteobacteria	1.00	CHR	NC_012034	C. bescii	Clostridia	0.82
RECE	NC_005140	V. vulnificus	γ proteobacteria	1.00	CHR				0.82
RECE	NC_014966					NC_014652	$C.\ hydrothermalis$	Clostridia	0.02
RECE		$V. \ vulnificus$	γ proteobacteria	1.00	CHR	NC_014652 NC_015949	$C. hydrothermalis \\ C. lactoaceticus$	Clostridia Clostridia	0.82
	NC_013457	V. vulnificus Vibrio sp.	γ proteobacteria γ proteobacteria	1.00 1.00	CHR CHR	NC_014652 NC_015949 NC_014721	C. hydrothermalis C. lactoaceticus C. kristjanssonii	Clostridia Clostridia Clostridia	0.82
RECE	NC_013457 NC_016614	V. vulnificus Vibrio sp. Vibrio sp.	$\gamma \mathrm{proteobacteria} \ \gamma \mathrm{proteobacteria} \ \gamma \mathrm{proteobacteria} \ \gamma \mathrm{proteobacteria}$	$1.00 \\ 1.00 \\ 0.36$	CHR CHR CHR	NC_014652 NC_015949 NC_014721 NC_014720	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis	Clostridia Clostridia Clostridia Clostridia	0.82 0.82 1.00 0.83
RECE	NC_013457 NC_016614	V. vulnificus Vibrio sp. Vibrio sp.	γ proteobacteria γ proteobacteria γ proteobacteria	$ \begin{array}{c} 1.00 \\ 1.00 \\ 0.36 \end{array} $	CHR CHR CHR CHR	NC_014652 NC_015949 NC_014721 NC_014720 NC_014392	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis	Clostridia Clostridia Clostridia Clostridia Clostridia	0.82 0.82 1.00 0.83 1.00
RECE	NC_013457 NC_016614 27 stability me	V. vulnificus Vibrio sp. Vibrio sp. asure : 0.769526	$\gamma \mathrm{proteobacteria} \ \gamma \mathrm{proteobacteria} \ \gamma \mathrm{proteobacteria} \ \gamma \mathrm{proteobacteria}$	$1.00 \\ 1.00 \\ 0.36$	CHR CHR CHR CHR CHR	NC_014652 NC_015949 NC_014721 NC_014720 NC_014392 NC_014657	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	0.82 1.00 0.83 1.00 0.50
RECE	NC_013457 NC_016614 27 stability me NC_006823	V. vulnificus Vibrio sp. Vibrio sp. asure : 0.769526 A. aromaticum	γ proteobacteria γ proteobacteria γ proteobacteria β proteobacteria	$ \begin{array}{c} 1.00 \\ 1.00 \\ 0.36 \end{array} $ 0.57	CHR CHR CHR CHR CHR CHR	NC_014652 NC_015949 NC_014721 NC_014720 NC_014392 NC_014657 NC_009437	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharaluticus	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.82 \\ 0.82 \\ 1.00 \\ 0.83 \\ 1.00 \\ 0.50 \\ 0.50 \end{array}$
RECE cluster : PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC 010553	V. vulnificus Vibrio sp. Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria	γproteobacteria γproteobacteria γproteobacteria βproteobacteria	$ \begin{array}{c} 1.00 \\ 1.00 \\ 0.36 \end{array} $	CHR CHR CHR CHR CHR CHR	NC_014652 NC_015949 NC_014721 NC_014720 NC_014392 NC_014657 NC_009437 NC_018264	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Terman eenhe sterium	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.82\\ 0.82\\ 1.00\\ 0.83\\ 1.00\\ 0.50\\ 0.50\\ 1.00\\ \end{array}$
RECE cluster : PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_008545	V. vulnificus Vibrio sp. Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocenacia	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array} $	CHR CHR CHR CHR CHR CHR phage	NC_014652 NC_015949 NC_014721 NC_014720 NC_014392 NC_014657 NC_009437 NC_018264	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.02\\ 0.82\\ 1.00\\ 0.83\\ 1.00\\ 0.50\\ 0.50\\ 1.00\\ \end{array}$
RECE cluster : PL PL PL PL PL	NC_013457 NC_016614 27 stability me. NC_006823 NC_010553 NC_008545 NC_011003	V. vulnificus Vibrio sp. Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array} $	CHR CHR CHR CHR CHR CHR phage	NC_014652 NC_015949 NC_014721 NC_014720 NC_014392 NC_014657 NC_009437 NC_018264	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	0.82 0.82 1.00 0.83 1.00 0.50 0.50 1.00
RECE cluster : PL PL PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_008545 NC_011003 NC_015227	V. vulnificus Vibrio sp. Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia B. cenocepacia	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array} $	CHR CHR CHR CHR CHR CHR phage	NC_014652 NC_015949 NC_014721 NC_014720 NC_014392 NC_014657 NC_009437 NC_018264	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	0.82 1.00 0.83 1.00 0.50 0.50 1.00
RECE cluster : PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_008545 NC_011003 NC_015377	V. vulnificus Vibrio sp. Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia B. gladioli D. gladioli	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 0.36 0.57 1.00 1.00 0.25 1.00	CHR CHR CHR CHR CHR phage	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014392 NC.014657 NC.009437 NC.009437 NC.018264	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage Masure : 0.824444	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.82\\ 0.82\\ 1.00\\ 0.83\\ 1.00\\ 0.50\\ 0.50\\ 1.00\\ \end{array}$
RECE cluster : PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_01537 NC_015377 NC_015378	V. vulnificus Vibrio sp. Asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia B. gladioli B. gladioli	γ proteobacteria γ proteobacteria γ proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array} $	CHR CHR CHR CHR CHR phage cluster CHR	NC_014652 NC_015949 NC_014721 NC_014720 NC_014720 NC_014657 NC_009437 NC_018264 : 32 stability m NC_014622	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage teasure : 0.824444 P. polymyza	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli	0.82 1.00 0.83 1.00 0.50 0.50 1.00
RECE PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me. NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015378 NC_015378	V. vulnificus Vibrio sp. vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia B. gladioli B. gladioli B. gladioli	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \\ 0.57\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$	CHR CHR CHR CHR CHR phage cluster CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014392 NC.014657 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014483	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. owensensis C. saccharolyticus Termoanaerobacterium phage measure : 0.824444 P. polymyxa P. polymyxa	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli	0.82 0.82 1.00 0.83 1.00 0.50 1.00 1.00 1.00
RECE PL PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me: NC_006823 NC_008545 NC_011003 NC_015377 NC_015378 NC_015383 NC_015233	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli	γ proteobacteria γ proteobacteria γ proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria β proteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array} $	CHR CHR CHR CHR CHR phage cluster CHR CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014392 NC.014392 NC.01457 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014833 NC.017542	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. krootskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage teasure : 0.824444 P. polymyza P. polymyza P. polymyza	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli	0.82 0.82 1.00 0.83 1.00 0.50 1.00 1.00 1.00 1.00
RECE PL PL PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010537 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \\ 0.57\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.67\\ 0.67\\ \hline \end{array}$	CHR CHR CHR CHR CHR CHR Phage Cluster CHR CHR CHR CHR CHR PL	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014392 NC_014392 NC_014392 NC_01437 NC_018264 : 32 stability m NC_014622 NC_014632 NC_017543	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage teasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli	0.82 0.82 1.00 0.83 1.00 0.50 1.00 1.00 1.00 1.00 1.00
RECE cluster : PL PL PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_015537 NC_015377 NC_015378 NC_015378 NC_015378 NC_012723 NC_012725 NC_012718	V. vulnificus Vibrio sp. Asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae	γ proteobacteria γ proteobacteria γ proteobacteria β proteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \\ 0.57\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.67\\ 1.00\\ \end{array}$	CHR CHR CHR CHR CHR phage CHR CHR CHR PL CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014657 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014622 NC.014483 NC.017542 NC.017543 NC.016641	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli	0.82 0.82 1.00 0.83 1.00 0.50 1.00 1.00 1.00 1.00 1.00
RECE PL PL PL PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me: NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015378 NC_015383 NC_012723 NC_012725 NC_012718 NC_012720	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR Phage cluster CHR CHR CHR PL CHR PL	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014392 NC.014657 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014622 NC.014483 NC.017542 NC.017543 NC.017543 NC.016641 NC.000959	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Deinococci	0.82 0.82 1.00 0.83 1.00 0.50 1.00 1.00 1.00 1.00 0.21
RECE cluster : PL PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010537 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012725 NC_012720 NC_010070	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. glumae	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR PL CHR PL	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014392 NC_014657 NC_009437 NC_018264 : 32 stability m NC_014622 NC_014622 NC_017542 NC_017543 NC_016641 NC_000959	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage teasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Deinococci	0.82 0.82 1.00 0.83 1.00 0.50 0.50 1.00 1.00 1.00 1.00 1.00 0.00 0.50
RECE cluster : PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010553 NC_015377 NC_015377 NC_015378 NC_01273 NC_012725 NC_012718 NC_012720 NC_010801	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. geladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR Phage CHR CHR CHR CHR PL CHR PL CHR PL	NC_014652 NC_014721 NC_014721 NC_014720 NC_014720 NC_014392 NC_014392 NC_014392 NC_014392 NC_01437 NC_018264 NC_018264 NC_017542 NC_016641 NC_000959	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage 16 16 17 17 18 19 19 10 19 10 10 10 10 10 10 10 10 10 10 10 10 10	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci	$\begin{array}{c} 0.82\\ 0.82\\ 1.00\\ 0.83\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.21\\ \end{array}$
RECE sluster : PL PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me: NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015378 NC_015378 NC_015383 NC_012723 NC_012725 NC_012720 NC_010070 NC_010801 NC_010802	V. vulnificus Vibrio sp. sasure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans	γρroteobacteria γρroteobacteria γρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria βρroteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR Phage cluster CHR CHR PL CHR PL cluster CUP	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014392 NC.014657 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014483 NC.017542 NC.017543 NC.017543 NC.016641 NC.000959 : 33 stability m	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. aanaperbases	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.82 0.82 1.00 0.83 1.00 0.50 1.00 1.00 1.00 1.00 0.21
RECE sluster : PL PL PL PL PL PL PL PL PL PL	NC_013457 NC_016614 27 stability me NC_006823 NC_008545 NC_011003 NC_015377 NC_015378 NC_015378 NC_012723 NC_012725 NC_012725 NC_012720 NC_01070 NC_010801 NC_010802 NC_010087	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans B. multivorans	γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR PL CHR PL CHR CHR CHR CHR CHR	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014720 NC_014392 NC_014392 NC_014392 NC_01437 NC_018264 : 32 stability m NC_017542 NC_017543 NC_016641 NC_00959 : 33 stability m NC_002946	C. hydrothermalis C. lactoaceticus C. kirstjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage teasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans teasure : 0.540825 N. gonorrhoeae	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Deinococci	0.82 1.00 0.83 1.00 0.50 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
RECE cluster : PL RECE* PL RECE* PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010553 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010627	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans B. multivorans B. multivorans	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \end{array}$	CHR CHR CHR CHR CHR phage cluster CHR CHR PL CHR PL cluster CHR CHR CHR CHR	NC_014652 NC_014721 NC_014721 NC_014720 NC_014720 NC_014392 NC_014392 NC_014392 NC_014392 NC_01437 NC_018264 NC_01642 NC_017543 NC_016641 NC_000959 : 33 stability m NC_002946 NC_011035	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage teasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans teasure : 0.540825 N. gonorrhoeae N. gonorrhoeae	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci β proteobacteria β proteobacteria	0.82 0.82 1.00 0.83 1.00 0.50 0.50 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me: NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015378 NC_015378 NC_015383 NC_012723 NC_012725 NC_012718 NC_012720 NC_010070 NC_010801 NC_010802 NC_010087 NC_010627	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. genocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans	γρroteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR Phage Cluster CHR CHR PL CHR PL CHR CHR CHR CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014657 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014622 NC.014483 NC.017542 NC.016641 NC.000959 : 33 stability m NC.002946 NC.011035 NC.017511	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Ba	0.82 1.00 0.83 1.00 0.50 0.50 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
RECE sluster : PL	NC_013457 NC_016614 27 stability me: NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015377 NC_015378 NC_015383 NC_012723 NC_012725 NC_012725 NC_012720 NC_0100801 NC_010801 NC_010802 NC_010627 NC_010627 NC_010627	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans	γproteobacteria γproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014720 NC_014720 NC_014720 NC_009437 NC_018264 : 32 stability m NC_014622 NC_017542 NC_017543 NC_017543 NC_017543 NC_00959 : 33 stability m NC_002946 NC_017511 NC_014752	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Deinococci βproteobacteria βproteobacteria βproteobacteria	0.82 0.82 1.00 0.83 1.00 0.50 0.50 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010553 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010627 NC_010627 NC_014718 NC_009227	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. vietnamiensis	γproteobacteria γproteobacteria γproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \end{array}$	CHR CHR CHR CHR CHR phage CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014720 NC_014657 NC_009437 NC_018264 : 32 stability m NC_014622 NC_014622 NC_014641 NC_000959 : 33 stability m NC_002946 NC_011035 NC_014751 NC_014752 NC_003112	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage teasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans teasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci βproteobacteria βproteobacteria βproteobacteria	0.82 0.82 1.00 0.50 0.50 0.50 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me: NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015378 NC_015378 NC_015383 NC_012723 NC_012725 NC_012718 NC_012720 NC_010801 NC_010801 NC_010802 NC_010087 NC_010627 NC_010227 NC_009230	V. vulnificus Vibrio sp. Asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. vietnamiensis B. vietnamiensis	γproteobacteria γproteobacteria γproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}\\ \hline \\ 0.57\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.67\\ 0.67\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.27\\ 1.00\\ 0.64\\ \end{array}$	CHR CHR CHR CHR CHR Phage Cluster CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014721 NC.014720 NC.014720 NC.014721 NC.014720 NC.014720 NC.014720 NC.01432 NC.018264 NC.014622 NC.014622 NC.014483 NC.017542 NC.016641 NC.000959 33 stability m NC.01035 NC.017511 NC.014752 NC.017512 NC.003112 NC.003116	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. saccharolyticus Termoanaerobacterium phage measure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans measure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. lactamica N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Deinococci β proteobacteria β proteobacteria β proteobacteria β proteobacteria	0.82 1.00 0.83 1.00 0.50 0.50 1.00
RECE sluster : PL	NC_013457 NC_016614 27 stability me: NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015377 NC_015383 NC_012723 NC_012723 NC_012725 NC_012720 NC_0100801 NC_010801 NC_010802 NC_010627 NC_010627 NC_010627 NC_010627 NC_010627 NC_010627 NC_010627 NC_009230 NC_014120	V. vulnificus Vibrio sp. A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. multivorans B. thizoxinica	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014720 NC_014720 NC_014720 NC_014720 NC_01437 NC_018264 32 stability m NC_014622 NC_017542 NC_017543 NC_017543 NC_017543 NC_007549 33 stability m NC_002946 NC_01035 NC_017511 NC_014752 NC_003112 NC_00316 NC_00316	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. ovensensis C. ovensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. lactamica N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli	0.82 0.82 1.00 0.83 1.00 0.50 0.50 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010537 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012725 NC_012727 NC_010801 NC_010801 NC_010802 NC_010627 NC_010627 NC_010627 NC_014718 NC_009227 NC_014120 NC_016591	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. genocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. multivorans B. vietnamiensis B. vietnamiensis B. vietnamiensis Burkholderia sp. Burkholderia sp.	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014720 NC_014657 NC_009437 NC_018264 : 32 stability m NC_014622 NC_014622 NC_017543 NC_017543 NC_016641 NC_000959 : 33 stability m NC_002946 NC_011035 NC_014752 NC_003112 NC_003116 NC_00316 NC_00316	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. saccharolyticus Termaanaerobacterium phage teasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans teasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. meningitidis N. meningitidis N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	0.82 0.82 1.00 0.50 0.50 0.50 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010553 NC_015377 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010801 NC_010802 NC_01087 NC_014718 NC_009227 NC_014718 NC_009227 NC_014718 NC_009230 NC_014718 NC_009230 NC_014718	V. vulnificus Vibrio sp. Asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis Burkholderia sp. C. metallidurans	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}\\ \hline \\ 0.57\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.67\\ 0.67\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.27\\ 1.00\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.12\\ 1.00\\ \end{array}$	CHR CHR CHR CHR CHR Phage CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014720 NC.014392 NC.014657 NC.009437 NC.018264 NC.018264 NC.014622 NC.014622 NC.014622 NC.014623 NC.017543 NC.016641 NC.000959 : 33 stability m NC.002946 NC.011035 NC.017511 NC.003112 NC.003116 NC.00316	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	0.82 0.00 1.00 0.50 0.50 0.50 1.00
RECE sluster : PL	NC_013457 NC_016614 27 stability me: NC_006823 NC_010553 NC_008545 NC_011003 NC_015377 NC_015378 NC_015378 NC_015383 NC_012723 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010087 NC_010627 NC_010627 NC_010627 NC_010627 NC_009230 NC_014120 NC_016591 NC_006525 NC_007971	V. vulnificus Vibrio sp. A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans C. metallidurans C. metallidurans	γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014392 NC.014657 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014683 NC.017542 NC.016641 NC.000959 : 33 stability m NC.017511 NC.014752 NC.011035 NC.017511 NC.014752 NC.003116 NC.00316 NC.010120 NC.017501	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. lactamica N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Deinococci βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	0.82 0.82 1.00 0.83 1.00 0.50 0.50 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010533 NC_015378 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010627 NC_010627 NC_010627 NC_010627 NC_014120 NC_014120 NC_016591 NC_007972	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans D. rhizoxinica B. vietnamiensis B. vietnamiensis Burkholderia sp. C. metallidurans C. metallidurans	γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR PL CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014652 NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014720 NC_009437 NC_009437 NC_018264 NC_014622 NC_014622 NC_017542 NC_017543 NC_016641 NC_000959 : 33 stability m NC_002946 NC_01035 NC_014752 NC_003116 NC_003116 NC_003116 NC_00120 NC_013016 NC_017551 NC_017551	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	0.82 0.82 1.00 0.50 0.50 0.50 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010553 NC_015377 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010687 NC_010687 NC_014718 NC_009227 NC_00476591 NC_006525 NC_007971 NC_006466	V. vulnificus Vibrio sp. Asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis C. metallidurans C. metallidurans	γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}\\ \hline \\ 0.57\\ 1.00\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.67\\ 0.67\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.27\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.12\\ 1.00\\ 0.83\\ 1.00\\ 0.83\\ 1.00\\ 0.00\\ 1.00\\ 1.00\\ 0.00$	CHR CHR CHR CHR CHR Phage CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014720 NC.014657 NC.009437 NC.018264 32 stability m NC.016622 NC.016641 NC.000959 33 stability m NC.002946 NC.017513 NC.017511 NC.003116 NC.00316 NC.013016 NC.017505 NC.017512	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. saccharolyticus Termoanaerobacterium phage 1000000000000000000000000000000000000	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci \$proteobacteria	0.82 0.00 1.00 0.50 0.50 0.50 1.00
RECE sluster : PL	NC_013457 NC_016614 27 stability me NC_006823 NC_008545 NC_011003 NC_015377 NC_015377 NC_015378 NC_015383 NC_012723 NC_012723 NC_012725 NC_012720 NC_010801 NC_010801 NC_010801 NC_010802 NC_010087 NC_010627 NC_010627 NC_010627 NC_009230 NC_014120 NC_016591 NC_007971 NC_007971 NC_007971 NC_007972 NC_00736	V. vulnificus Vibrio sp. Asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans C. metallidurans C. metallidurans C. metallidurans	γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \hline \end{array}$	CHR CHR CHR CHR CHR CHR CHR CHR CHR PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014652 NC.014721 NC.014720 NC.014720 NC.014720 NC.014657 NC.009437 NC.018264 : 32 stability m NC.014622 NC.014622 NC.014622 NC.014483 NC.017543 NC.016641 NC.000959 : 33 stability m NC.002946 NC.011035 NC.017511 NC.003116 NC.003116 NC.00316 NC.017501 NC.017505 NC.017515 NC.017512	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. lactamica N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Bacilli Deinococci \$proteobacteria	1.00 0.82 1.00 0.83 1.00 0.50 0.50 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
RECE cluster : c	NC_013457 NC_016614 27 stability me NC_006823 NC_010533 NC_008545 NC_011003 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012725 NC_012725 NC_012725 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010627 NC_010627 NC_010627 NC_014120 NC_016591 NC_007972 NC_007972 NC_006466 NC_007336	V. vulnificus Vibrio sp. asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans C. metallidurans C. metallidurans	γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}$	CHR CHR CHR CHR CHR Phage CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_014652 NC_014721 NC_014720 NC_014720 NC_014720 NC_014720 NC_014720 NC_014657 NC_009437 NC_018264 : 32 stability m NC_014622 NC_017542 NC_017543 NC_017543 NC_017543 NC_0107511 NC_002946 NC_010135 NC_014752 NC_003112 NC_003116 NC_00316 NC_017501 NC_017505 NC_017512 NC_017513	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. owensensis C. saccharolyticus Termaanaerobacterium phage teasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans teasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. lactamica N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	0.82 0.82 1.00 0.50 0.50 0.50 1.00
RECE :luster : PL	NC_013457 NC_016614 27 stability me NC_006823 NC_010553 NC_010553 NC_015377 NC_015377 NC_015378 NC_012723 NC_012725 NC_012725 NC_012720 NC_010801 NC_010801 NC_010802 NC_010627 NC_010627 NC_010627 NC_010627 NC_016591 NC_006525 NC_007971 NC_006466 NC_007336 NC_015724	V. vulnificus Vibrio sp. Asure : 0.769526 A. aromaticum B. ambifaria B. cenocepacia B. gladioli B. gladioli B. gladioli B. gladioli B. gladioli B. glumae B. glumae B. glumae B. glumae B. glumae B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. multivorans B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis B. vietnamiensis C. metallidurans C. metallidurans C. metallidurans C. metallidurans C. metallidurans C. metallidurans C. metallidurans C. metallidurans C. metallidurans	γproteobacteria γproteobacteria γproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.36\\ \end{array}\\ \hline \\ 0.57\\ 1.00\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.67\\ 0.67\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.73\\ 0.45\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.64\\ 1.00\\ 0.83\\ 1.00\\ 0.83\\ 1.00\\ 0.83\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.0\\ 0.$	CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC.014652 NC.015949 NC.014721 NC.014720 NC.014720 NC.014720 NC.014392 NC.014657 NC.009437 NC.018264 NC.018264 NC.014622 NC.014622 NC.017543 NC.017543 NC.016641 NC.000959 : 33 stability m NC.002946 NC.017513 NC.013016 NC.003116 NC.0017501 NC.017501 NC.017513 NC.017513 NC.017513	C. hydrothermalis C. lactoaceticus C. kristjanssonii C. kronotskyensis C. obsidiansis C. obsidiansis C. saccharolyticus Termoanaerobacterium phage neasure : 0.824444 P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. polymyxa P. terrae D. radiodurans neasure : 0.540825 N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. gonorrhoeae N. lactamica N. meningitidis N. meningitidis	Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Bacilli Bacilli Bacilli Bacilli Deinococci \$proteobacteria	0.82 1.00 0.83 1.00 0.50 0.50 1.00

263
Annexe D.Résultats du clustering de V^R par INFOMAP

arre				1 4 6 6		229 040000			
CHR	NC_017516	N. meningitidis	proteobacteria	1.00	PL	NC_010893	M. bovis	γ proteobacteria	1.00
CHR	NC_017517	N. meningitidis	proteobacteria	1.00	PL	NC_010900	M. bouis	γ proteobacteria	1.00
DI	NC 012518	N. meningulais	^p proteobacteria	1.00		NC 011121	M. oovis M. aatarmhalia	γproteobacteria	1.00
FL	NC-015518	5. termitiais	rusobacterna	0.27	PL PI	NC 003411	M. catarrhaus M. variaena	γproteobacteria	1.00
cluster	: 34 stability	measure : 0.985000			PL	NC 004772	P. multocida	vproteobacteria	1.00
CHB	NC 012960	Cand Hodgkinia	oproteobacteria	0.97	PL	NC_005921	P. suringae	vproteobacteria	1.00
CHR	NC_010842	L. biflexa	Spirochaetes	1.00	\mathbf{PL}	NC_019134	S. enterica	γ proteobacteria	1.00
CHR	NC_010602	L. biflexa	Spirochaetes	1.00	\mathbf{PL}	NC_008320	Shewanella sp.	γ proteobacteria	1.00
CHR	NC_008508	L. borgpetersenii	Spirochaetes	1.00	\mathbf{PL}	NC_002144	Y. pestis	γ proteobacteria	0.36
CHR	NC_008510	L. borgpetersenii	Spirochaetes	1.00					
CHR	NC_004342	L. interrogans	Spirochaetes	1.00	cluster	: 37 stability	measure : 0.734815		
CHR	NC_005823	L. interrogans	Spirochaetes	1.00	CHR	NC_017030	C. coralloides	δ proteobacteria	1.00
CHR	NC_017551	L. interrogans	Spirochaetes	1.00	CHR	NC_013440	H. ochraceum	δ proteobacteria	0.82
CHR	NC_018020	T. parva	Spirochaetes	1.00	CHR	NC_015711	M. fulvus	δ proteobacteria	1.00
					CHR	NC_008095	M. xanthus	δ proteobacteria	1.00
cluster	: 35 stability	measure : 1.000000	<u> </u>	1.00	CHR	NC_014623	S. aurantiaca	∂proteobacteria	1.00
CHR	NC_008277	B. afzelii	Spirochaetes	1.00	• •	00 1 1 11	0.050104		
CHR	NC_017238	B. afzelii	Spirochaetes	1.00	cluster	: 38 stability	measure : 0.858104		1.00
CHR	NC_015021	B. afzeln B. hissouthi	Spirochaetes	1.00	PL	NC_011732	Cyanothece sp.	Chroobacteria	1.00
CHR	NC 001218	B. Dissettiii B. humadomfori	Spirochaetes	1.00	PL	NC 012167	Cyanothece sp.	Chroobacteria	1.00
CHR	NC 011728	B. burgdorferi B. burgdorferi	Spirochaetes	1.00	PL	NC 019429	Anahaena sp.	Homogoneae	1.00
CHR	NC 017403	B. burgdorferi	Spirochaetes	1.00	PL	NC 012885	A hudrophila	oproteobacteria	1.00
CHR	NC 017418	B. buradorferi	Spirochaetes	1.00	PL	NC 009349	A. salmonicida	vproteobacteria	1.00
CHR	NC_017808	B. crocidurae	Spirochaetes	1.00	PL	NC_012886	E. coli	γ proteobacteria	1.00
CHR	NC_011229	B. duttonii	Spirochaetes	1.00	\mathbf{PL}	NC_012690	E. coli	γ proteobacteria	nan
CHR	NC_006156	B. garinii	Spirochaetes	nan	\mathbf{PL}	NC_012692	E. coli	γ proteobacteria	1.00
CHR	NC_017717	B. garinii	Spirochaetes	1.00	\mathbf{PL}	NC_017645	E. coli	γ proteobacteria	1.00
CHR	NC_018747	B. garinii	Spirochaetes	1.00	\mathbf{PL}	NC_018994	E. coli	γ proteobacteria	1.00
CHR	NC_010673	B. hermsii	Spirochaetes	1.00	PL	NC_019045	E. coli	γ proteobacteria	1.00
CHR	NC_011244	B. recurrentis	Spirochaetes	1.00	PL	NC_019065	E. coli	γ proteobacteria	1.00
CHR	NC_008710	B. turicatae	Spirochaetes	1.00	PL	NC_019066	E. coli	γ proteobacteria	1.00
					PL	NC_019069	E. coli	γ proteobacteria	1.00
cluster	: 36 stability	measure : 0.859439			PL	NC_016839	K. pneumoniae	γ proteobacteria	1.00
PL	NC_012439	P. marina	Aquificae	0.50	PL	NC_016976	K. pneumoniae	γ proteobacteria	1.00
PL	NC_014031	B. megaterium	Bacilli	1.00	PL	NC_019153	K. pneumoniae	γ proteobacteria	1.00
PL	NC_013540	Planococcus sp.	Cleateridia	1.00		NC 008612	R. pneumoniae	γproteobacteria	1.00
PL	NC 011370	B. fibrisolvens	Clostridia	1.00	PL.	NC 008613	P damselae	oproteobacteria	1.00
PL	NC 017998	Thermoan aerobacterium	Clostridia	0.00	PL	NC 016983	P damselae	oproteobacteria	1.00
	1101011000	phage	Clostildia	0.00	PL	NC_009444	P. fluorescens	γ proteobacteria	1.00
PL	NC_018745	E. oligotrophica	Cvtophagia	1.00	\mathbf{PL}	NC_016974	P. stuartii	γ proteobacteria	1.00
PL	NC_018744	E. oligotrophica	Cytophagia	1.00	\mathbf{PL}	NC_003905	P. vulgaris	γ proteobacteria	1.00
PL	NC_019017	F. limi	Cytophagia	0.41	\mathbf{PL}	NC_009140	S. enterica	γ proteobacteria	1.00
PL	NC_015693	R. slithyformis	Cytophagia	1.00	\mathbf{PL}	NC_012693	S.~enterica	γ proteobacteria	1.00
PL	NC_015694	R. slithyformis	Cytophagia	1.00	\mathbf{PL}	NC_019107	S.~enterica	γ proteobacteria	1.00
PL	NC_013732	S. linguale	Cytophagia	0.83	PL	NC_019116	$S. \ enterica$	γ proteobacteria	1.00
PL	NC_013734	S. linguale	Cytophagia	1.00	\mathbf{PL}	NC_019118	$S. \ enterica$	γ proteobacteria	1.00
PL	NC_013735	S. linguale	Cytophagia	1.00	PL	NC_019121	S. enterica	γ proteobacteria	1.00
PL	NC_013736	S. linguale	Cytophagia	1.00	PL	NC_005250	V. anguillarum	γ proteobacteria	1.00
PL	NC_014157	S. ruber	Cytophagia	1.00	PL	NC_011185	V. fischeri	γ proteobacteria	1.00
PL	NC_018751	F. branchiophilum	Flavobacterna	1.00	PL	NC_014170	A. nematopnua V. martin	γ proteobacteria	1.00
PL	NC 012033	n. nyurossis Pedobacter sp	Sphingobacteriia	1.00	PL.	NC 009139	1. pestis V ruckeri	oproteobacteria	1.00
PL	NC 017071	S suminantium	Norativicutos	1.00	11	110_009139	1. 1000011	yproteobacteria	1.00
PL	NC 017923	Burkholderia sp	Bproteobacteria	1.00	cluster	: 39 stability	measure : 0.624229		
PL	NC_010881	N. gonorrhoeae	βproteobacteria	1.00	CHR	NC_013861	L. lonabeachae	γproteobacteria	1.00
PL	NC_010855	N. lactamica	β proteobacteria	1.00	CHR	NC_002942	L. pneumophila	γ proteobacteria	1.00
PL	NC_010871	N. lactamica	β proteobacteria	1.00	CHR	NC_006368	L. pneumophila	γ proteobacteria	1.00
PL	NC_010888	N. lactamica	β proteobacteria	1.00	CHR	NC_006369	L. pneumophila	γ proteobacteria	1.00
PL	NC_010906	N. lactamica	β proteobacteria	1.00	CHR	NC_009494	L. pneumophila	γ proteobacteria	1.00
PL	NC_010928	N. lactamica	β proteobacteria	1.00	CHR	NC_016811	L. pneumophila	γ proteobacteria	1.00
PL	NC_010683	R. pickettii	β proteobacteria	1.00	CHR	NC_018139	L. pneumophila	γ proteobacteria	1.00
PL	NC_012851	R. pickettii	β proteobacteria	1.00	CHR	NC_018140	L. pneumophila	γ proteobacteria	1.00
гL рт	NC 000706	D. marinus	oproteobacteria	0.70	онк БТ	NC 006960	L. pneumopnila	γproteobacteria	1.00
PL	NC 006877	A baumannii	oproteobacteria	1.00	гь	110-000800	r. choterue	/ proteobacteria	0.04
PI.	NC 009083	A. baumannii	vproteobacteria	1.00	cluster	: 40 stability	measure · 0 506151		
PL	NC_009084	A. baumannii	γproteobacteria	1.00	PL	NC.011311	A. salmonicida	γproteobacteria	0.86
PL	NC_010395	A. baumannii	γ proteobacteria	1.00	PL	NC_002525	E. coli	γ proteobacteria	0.00
PL	NC_010396	A. baumannii	γ proteobacteria	1.00	\mathbf{PL}	NC_005923	E. coli	γ proteobacteria	0.00
PL	NC_010398	A. baumannii	γ proteobacteria	1.00	\mathbf{PL}	NC_010257	E. coli	γ proteobacteria	0.00
PL	NC_010401	A. baumannii	γ proteobacteria	1.00	\mathbf{PL}	NC_013503	E. coli	γ proteobacteria	0.73
PL	NC_010402	A. baumannii	γ proteobacteria	1.00	PL	NC_015472	E. coli	γ proteobacteria	0.50
PL	NC_010481	A. baumannii	γ proteobacteria	1.00	PL	NC_019046	E. coli	γ proteobacteria	0.38
PL	NC_010605	A. baumannii	γ proteobacteria	1.00	PL	NC_019047	E. coli	γ proteobacteria	1.00
PL	NC_011585	A. baumannii	γ proteobacteria	1.00	PL	NC_019067	E. coli	γ proteobacteria	0.67
PL DT	NC 012506	A. baumannii	γproteobacteria	1.00		NC_019083	E. coli E. coli	γ proteobacteria	1.00
PL PI	NC 017165	A. baumannii	vproteobacteria	1.00		NC 010012	E. coli	γproteobacteria	0.00
PL	NC 017179	A baymannii	oproteobacteria	1.00	гь рт	NC 010020	E. coli	proteobacteria	1.00
PI.	110-01/11/2	A. calcoaceticus	vproteobacteria	1.00	PI.	NC 019039	E coli	vproteobacteria	1.00
	NC 013056		, processactoria	1.00	PL	NC.010697	E. tasmaniensis	vproteobacteria	0.57
PL 1	NC_013056 NC 005245	A. paragallingrum	vproteobacteria	1 1.000	1 11	110-010097		/ proteopacteria	0.07
PL PL	NC_013056 NC_005245 NC_011316	A. paragallinarum A. salmonicida	γ proteobacteria γ proteobacteria	1.00	PL	NC.019162	K. pneumoniae	vproteobacteria	0.57
PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309	A. paragallinarum A. salmonicida A. venetianus	γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 0.79	PL PL	NC_019162 NC_019157	K. pneumoniae K. pneumoniae	γ proteobacteria γ proteobacteria	$0.57 \\ 1.00$
PL PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309 NC_010310	A. paragallinarum A. salmonicida A. venetianus A. venetianus	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 0.79 1.00	PL PL PL	NC_019162 NC_019157 NC_019161	K. pneumoniae K. pneumoniae K. pneumoniae	γ proteobacteria γ proteobacteria γ proteobacteria	$0.57 \\ 1.00 \\ 1.00$
PL PL PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309 NC_010310 NC_008487	A. paragallinarum A. salmonicida A. venetianus A. venetianus E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 0.79 1.00 1.00	PL PL PL PL	NC_019162 NC_019157 NC_019161 NC_019384	K. pneumoniae K. pneumoniae K. pneumoniae K. pneumoniae	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$0.57 \\ 1.00 \\ 1.00 \\ 0.67$
PL PL PL PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309 NC_010310 NC_008487 NC_010486	A. paragallinarum A. salmonicida A. venetianus A. venetianus E. coli E. coli	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.79\\ 1.00\\ 1.00\\ 1.00 \end{array} $	PL PL PL PL PL	NC_019162 NC_019157 NC_019161 NC_019384 NC_016036	K. pneumoniae K. pneumoniae K. pneumoniae K. pneumoniae M. morganii	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$0.57 \\ 1.00 \\ 1.00 \\ 0.67 \\ 0.64$
PL PL PL PL PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309 NC_010310 NC_008487 NC_010486 NC_010885	A. paragallinarum A. salmonicida A. venetianus A. venetianus E. coli E. coli E. coli	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.79\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array} $	PL PL PL PL PL PL	NC_019162 NC_019157 NC_019161 NC_019384 NC_016036 NC_007968	K. pneumoniae K. pneumoniae K. pneumoniae K. pneumoniae M. morganii P. cryohalolentis	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	0.57 1.00 1.00 0.67 0.64 1.00
PL PL PL PL PL PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309 NC_010310 NC_008487 NC_010486 NC_010885 NC_011411	A. paragallinarum A. salmonicida A. venetianus E. coli E. coli E. coli E. coli	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$ \begin{array}{c} 1.00\\ 1.00\\ 0.79\\ 1.00$	PL PL PL PL PL PL PL	NC_019162 NC_019157 NC_019161 NC_019384 NC_016036 NC_007968 NC_014653	K. pneumoniae K. pneumoniae K. pneumoniae M. morganii P. cryohalolentis P. damselae	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	0.57 1.00 1.00 0.67 0.64 1.00 0.68
PL PL PL PL PL PL PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309 NC_010310 NC_008487 NC_010486 NC_010486 NC_010485 NC_011411 NC_015461	A. paragallinarum A. salmonicida A. venetianus A. venetianus E. coli E. coli E. coli E. coli G. anatis	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.79\\ 1.00\\$	PL PL PL PL PL PL PL	NC_019162 NC_019157 NC_019161 NC_019384 NC_016036 NC_007968 NC_007968 NC_014653 NC_005871	K. pneumoniae K. pneumoniae K. pneumoniae M. morganii P. cryohalolentis P. damselae P. profundum	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	0.57 1.00 1.00 0.67 0.64 1.00 0.68 1.00
PL PL PL PL PL PL PL PL PL PL PL	NC_013056 NC_005245 NC_011316 NC_010309 NC_010310 NC_008487 NC_010486 NC_010885 NC_011411 NC_015461 NC_006298	A. paragallinarum A. salmonicida A. venetianus E. coli E. coli E. coli E. coli G. anatis H. somnus M. sloci i coli	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 0.79\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.72\end{array}$	PL PL PL PL PL PL PL PL PL	NC_019162 NC_019157 NC_019161 NC_019384 NC_016036 NC_007968 NC_007968 NC_007967 NC_005871 NC_010555	K. pneumoniae K. pneumoniae K. pneumoniae M. morganii P. cryohalolentis P. damselae P. profundum P. mirabilis	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.57 \\ 1.00 \\ 1.00 \\ 0.67 \\ 0.64 \\ 1.00 \\ 0.68 \\ 1.00 \\ 0.50 \end{array}$

\mathbf{PL}	NC_015708	Pseudoalteromonas sp.	γ proteobacteria	1.00	PL	NC_018633	C. psittaci	Chlamydiia	1.00
\mathbf{PL}	NC_010860	S. enterica	γ proteobacteria	0.88	\mathbf{PL}	NC_018634	C. psittaci	Chlamydiia	1.00
PL.	NC 010421	S enterica	oproteobacteria	0.50	PL.	NC 018635	C. psittaci	Chlamydija	1 00
1 1	100_010421	D. enterica	proteobacteria	0.50		NC_010033	C. psilluci	Chianiyuna	1.00
PL	NC_019129	S. enterica	γ proteobacteria	0.00	PL	NC_018636	C. psittaci	Chlamydiia	1.00
PL	NC_019138	S. enterica	γ proteobacteria	0.50	PL	NC_018637	C. psittaci	Chlamydiia	1.00
PL	NC 009036	S baltica	vproteobacteria	0.00	PL	NC 018638	C nsittaci	Chlamydija	1 00
DI	NC 000661	C L_14:	, provide a la stania	0.25	DI	NC 018620	C	Chlamodiia	1.00
FL	NC_009001	S. ballica	proteobacteria	0.25	FL	NC_018039	C. psillaci	Chianyuna	1.00
PL	NC_009998	S. baltica	γ proteobacteria	0.88	PL	NC_018640	C. psittaci	Chlamydiia	1.00
\mathbf{PL}	NC_011664	S. baltica	γ proteobacteria	0.77	PL	NC_019392	C. psittaci	Chlamydiia	1.00
PL.	NC 011668	S baltica	oproteobacteria	0.50	PL.	NC 002182	C muridarum	Chlamydija	1.00
 DI	NG 010005	a Lu:	/protocobactoria	1.00		NG 001270		Cl la l'in	1.00
PL	NC-016905	S. baltica	γ proteobacteria	1.00	PL	NC_001372	C. trachomatis	Chlamydila	1.00
\mathbf{PL}	NC_017578	S. baltica	γ proteobacteria	0.00	\mathbf{PL}	NC_007430	C. trachomatis	Chlamydiia	1.00
\mathbf{PL}	NC_017570	S. baltica	γ proteobacteria	0.71	PL	NC_010029	C. trachomatis	Chlamydija	1.00
DI	NC 017579	S halting	* protochastoria	1.00	DI	NC 010285	C trachomatic	Chlamydiin	1.00
F L	NC_017572	S. bannea	yproteobacteria	1.00	F L	NC_010285	C. trachomatis	Cinaliyuna	1.00
PL	NC_004349	S. oneidensis	γ proteobacteria	1.00	PL	NC_010286	C. trachomatis	Chlamydiia	1.00
\mathbf{PL}	NC_008573	Shewanella sp.	γ proteobacteria	0.00	PL	NC_012625	C. trachomatis	Chlamydiia	1.00
PL.	NC 012035	Shewanella sp	oproteobacteria	0.36	PL.	NC 012626	C. trachomatis	Chlamydija	1.00
DI	NG 010000	V I III	proteobacteria	1.00		NC 010007		Clib	1.00
PL	NC_012209	Y. enterocolítica	γ proteobacteria	1.00	PL	NC_012627	C. trachomatis	Chlamydila	1.00
\mathbf{PL}	NC_015054	Y. pestis	γ proteobacteria	0.36	PL	NC_012629	C. trachomatis	Chlamydiia	nan
					PL	NC_012630	C. trachomatis	Chlamvdiia	1.00
cluster	• 41 stability	$m_{easure} : 0.960968$			PL.	NC 012631	C. trachomatis	Chlamydija	1.00
GUD	NG contra	G II I I III	CILL LI	1.00		NG 017499		Cl la l'in	1.00
CHR	NC_007514	C. chlorochromatii	Chlorobia	1.00	PL	NC_017433	C. trachomatis	Chlamydila	1.00
\mathbf{CHR}	NC_010803	C. limicola	Chlorobia	1.00	PL	NC_017435	C. trachomatis	Chlamydiia	1.00
CHR	$NC_{-}007512$	C. luteolum	Chlorobia	1.00	PL	NC_017438	C. trachomatis	Chlamvdiia	1.00
CUP	NC 011027	C mammum	Chlorobia	1.00	PI	NC 015710	S negenereis	Chlamydiia	1.00
CIII	NC_011027	C. parvam	Ciliorobia	1.00		NG_010000	D. negevensis	Chianiyuna	1.00
CHR	NC_010831	C. phaeobacteroides	Chlorobia	1.00	PL	NC_014226	W. chondrophila	Chlamydiia	1.00
CHR	NC_008639	C. phaeobacteroides	Chlorobia	1.00					
CHB	NC 009337	C. phaeovibrioides	Chlorobia	1.00	cluster	· 44 stability	measure : 0.602222		
CUD	NG 000000	C. phaceboontonacs	Clibert	1.00	CUD	NG 011021		C111	1.00
CHR	NC_002932	C. tepiaum	Chlorobia	1.00	CHR	NC_011831	C. aggregans	Chloronexi	1.00
\mathbf{CHR}	NC_011059	P. aestuarii	Chlorobia	1.00	CHR	NC_010175	C. aurantiacus	Chloroflexi	1.00
CHR	NC_011060	P. phaeoclathratiforme	Chlorobia	1.00	CHR	NC_012032	Chloroflexus sp.	Chloroflexi	1.00
					CHD	NC 000767	D	Chlander	
• .					CIIR	NC-009707	n. castennoizn	Chiloronexi	nan
cluster	•:42 stability	measure : 0.949217			CHR	NC_009523	Roseiflexus sp.	Chloroflexi	1.00
PL	NC_016615	A. rhombi	Actinobacteridae	0.50					
PL.	NC 007068	B catenulatum	Actinobacteridae	1.00	cluster	· 45 stability	measure : 0 536399		
DI	NG 000500	D. Calchalatam	Autolacteridae	1.00	DI	NG 015060		A	0.64
PL	NC_002522	B. linens	Actinobacteridae	1.00	PL	NC_015060	G. tunaricola	Acidobacteria	0.64
\mathbf{PL}	NC_004253	B. longum	Actinobacteridae	1.00	\mathbf{PL}	NC_015560	A. subflavus	Actinobacteridae	0.67
PL	NC_011139	B. longum	Actinobacteridae	1.00	PL	NC_014030	S. ruber	Cvtophagia	0.71
DI	NC 011020	C agagi	Actinobactoridae	1.00	DI	NC 012212	1 mantouriamaun	oprotochastoria	1.00
FL	NC_011030	C. cusei	Actinobacteridae	1.00	FL	NC_015212	A. pasteurianas	aproteobacteria	1.00
PL	NC_001791	C. glutamicum	Actinobacteridae	1.00	PL	NC_017110	A. pasteurianus	α proteobacteria	1.00
PL	NC_002099	C. glutamicum	Actinobacteridae	1.00	PL	NC_017114	A. pasteurianus	α proteobacteria	1.00
\mathbf{PL}	NC 002611	P. acidipropionici	Actinobacteridae	1.00	PL	NC 017119	A. pasteurianus	oproteobacteria	1.00
DI	NC 002580	D fam. Jammi - Lii			DI	NC 017197	A		1.00
FL	NC_002580	F. freudenreichti	Actinobacteridae	nan	PL	NC_01/12/	A. pasteurianus	αproteobacteria	1.00
PL	NC_005705	P. jensenii	Actinobacteridae	0.93	PL	NC_017135	A. pasteurianus	α proteobacteria	1.00
\mathbf{PL}	NC_003846	R. erythropolis	Actinobacteridae	1.00	\mathbf{PL}	NC_017147	A. pasteurianus	α proteobacteria	1.00
PL.	NC 007487	R eruthropolis	Actinobacteridae	1.00	PL.	NC 017151	A pasteurianus	oproteobacteria	1 00
	NG 000000	D. D	A controbacteridae	1.00	DI	NG 000071	a l	aproteobacteria	1.00
PL	NC_006969	R. opacus	Actinobacteridae	1.00	PL	NC_006674	G. oxydans	α proteobacteria	1.00
PL	NC_008823	R. rhodochrous	Actinobacteridae	0.43	PL	NC_015059	G. tundricola	Acidobacteria	0.33
\mathbf{PL}	NC 004900	Rhodococcus sp.	Actinobacteridae	1.00	PL	NC 015058	G. tundricola	Acidobacteria	0.30
DI	NC 002142	C tootootoo op	Questo a bastania	1.00	DI	NC 019807	M	ttt	0.77
FL	NC_002145	C. lesiosleroni	pproteobacteria	1.00	PL	NC_012807	M. exiorquens	αproteobacteria	0.77
\mathbf{PL}	NC_010931	N. lactamica	β proteobacteria	0.64	PL	NC_010517	M. radiotolerans	α proteobacteria	0.63
PL	NC_006968	N. lactamica	β proteobacteria	1.00	PL	NC_010373	Methylobacterium sp.	α proteobacteria	0.21
PL	NC 010867	N lactamica	Bprotechacteria	1.00	PI	NC 003374	G oradans	oproteobacteria	0.54
	NG 010000	N. I. C.	pproteobacteria	1.00	DI	NG 010045	D . I'l	aproteobacteria	0.04
PL	NC_010926	N. lactamica	Bproteobacteria	1.00	PL	NC_010847	P. aminophilus	α proteobacteria	0.60
\mathbf{PL}	NC_016852	A. hydrophila	γ proteobacteria	1.00	\mathbf{PL}	NC_013788	Z. mobilis	α proteobacteria	0.21
\mathbf{PL}	NC_004923	A. salmonicida	γ proteobacteria	1.00	PL	NC_017183	Z. mobilis	αproteobacteria	0.00
DI	NC 004220	A calmonicida	* protochastoria	1.00	DI	NC 018147	Z mobilia	oprotochastoria	0.21
F L	NC_004339	A. suimoniciuu	yproteobacteria	1.00	F L	100_018147	2. moonis	aproteobacteria	0.21
PL	NC_004924	A. salmonicida	γ proteobacteria	0.95	PL	NC_016851	A. fischeri	γ proteobacteria	0.41
\mathbf{PL}	NC_004338	A. salmonicida	γ proteobacteria	1.00	PL	NC_006366	L. pneumophila	γ proteobacteria	0.31
PL	NC 013719	C rodentium	vproteobacteria	1.00	PL	NC 007186	S alossinidius	vproteobacteria	0.54
DI	NC 008480	Eli	/protocobactoria	1.00		110=001100	5. grocomatae	/protoobacteria	1 0.01
FL	NC_008489	E. coli	vproteobacteria	1.00					
PL	NC_009781	E. coli	γ proteobacteria	1.00	cluster	: 46 stability	measure : 0.878571		
\mathbf{PL}	NC_010883	E. coli	γ proteobacteria	1.00	PL	NC_016748	M. piezophila	Thermotogae	1.00
\mathbf{PL}	NC 011977	E. coli	vproteobacteria	1.00	CHB	NC 009828	T. lettingae	Thermotogae	1 00
DI	NC 010890	E1:	, provide a characteria	1.00	CIID	NG 000050	The internet and the internet in the internet internet in the internet intern	Thermotogue	1.00
	NG 012002	D. COM	proceonacteria	1.00	CHR	110_000803	1. manitima	i nermotogae	1.00
PL	NC_013367	E. coli	γ proteobacteria	1.00	CHR	NC_013642	1. naphthophila	Thermotogae	1.00
\mathbf{PL}	NC_017648	E. coli	γ proteobacteria	1.00	CHR	NC_011978	T. neapolitana	Thermotogae	1.00
\mathbf{PL}	NC 017723	E. coli	γproteobacteria	1.00	CHB	NC 009486	T. petrophila	Thermotogae	1 00
 DI	NG 010060	E l'	/protocobactoria	1.00	CIIIC	NG 015505	T. petrophila	Thermotogae	1.00
	110-019062	E. con	proceonacteria	1.00	UHR	INC_015707	1. inermarum	1 nermotogae	1.00
РL	NC_002498	E. ictaluri	γ proteobacteria	1.00	CHR	NC_010483	Thermotoga sp.	Thermotogae	1.00
\mathbf{PL}	NC_019160	K. pneumoniae	γ proteobacteria	1.00					
PL	NC 016840	K. pneumoniae	vproteobacteria	1.00	cluster	· 47 stabilit.	measure : 0.877045		
	NG 000700	V1-L-:-11.	/Pistobacteria	1.00	- cruster	stability	4 1		1.0.
PL	INC_003789	nieosiella sp.	γ proteobacteria	1.00	PL	NC_016597	A. brasilense	α proteobacteria	1.00
\mathbf{PL}	NC_016746	Oceanimonas sp.	γ proteobacteria	1.00	\mathbf{PL}	NC_009468	A. cryptum	α proteobacteria	0.10
\mathbf{PL}	NC_005325	M. catarrhalis	γ proteobacteria	1.00	PI.	NC_016624	A. lipoferum	aproteobacteria	0.73
PL	NC 006868	P multocida	vproteobactoria	0.64	л 2 DT	NC 012010	A nastaurianu-	oprotochesteria	1 00
 DT	NC 010275	D J 12	/proteobacteria	1.00	гь 	NG_013210	A. pusicurianus	aproteobacteria	1.00
гь –	110_010675	i seuuouiteromonas sp.	γproteopacteria	1.00	PL	NC_013211	A. pasteurianus	α proteobacteria	1.00
\mathbf{PL}	NC_009516	Psychrobacter sp.	γ proteobacteria	1.00	\mathbf{PL}	NC_017101	A. pasteurianus	α proteobacteria	1.00
\mathbf{PL}	NC_010659	S. boydii	γ proteobacteria	1.00	PI.	NC 017104	A. pasteurianus	aproteobacteria	1.00
PL	NC 014254	S enterica	oproteobactoria	1.00		NC 017105	A manter	operate the state of the	1.00
1 L D1	NG 011002		proceonacteria	1.00	PL 	NG_01/105	A. pasieurianus	aproteopacteria	1.00
PL	NC_011093	5. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_017113	A. pasteurianus	α proteobacteria	1.00
\mathbf{PL}	NC_003455	S. enterica	γ proteobacteria	1.00	\mathbf{PL}	NC_017117	A. pasteurianus	α proteobacteria	1.00
\mathbf{PL}	NC_010464	S. maltophilia	γproteobacteria	1.00	PI.	NC 017118	A pasteurianus	oproteobacteria	1 00
	NC 011707	V Auniali-	/Protobacteria	1.00	г L DI	NG 017110	A pusicultutus	aproteobacteria	1.00
PL D-	INC_011797	v. fluvialis	γ proteobacteria	1.00	PL	NC_017122	A. pasteurianus	α proteobacteria	1.00
PL	NC_012208	Y. enterocolitica	γ proteobacteria	1.00	\mathbf{PL}	NC_017126	A. pasteurianus	α proteobacteria	1.00
					\mathbf{PL}	NC_017130	A. pasteurianus	α proteobacteria	nan
cluster	: 43 stability	measure : 0.973591			PT	NC 017131	Anasteurianus	oprotechacteria	1 00
DT	NG oo 4700		C111	1.00	гь 57	NG_017131	A. pusicurianus	aproteobacteria	1.00
РL	NC_004720	C. caviae	Chlamydiia	1.00	\mathbf{PL}	NC_017134	A. pasteurianus	α proteobacteria	1.00
PL	NC_007900	C. felis	Chlamydiia	1.00	\mathbf{PL}	NC_017136	A. pasteurianus	α proteobacteria	1.00
\mathbf{PL}	NC_017286	C. pneumoniae	Chlamvdiia	1.00	PI.	NC 017143	A. pasteurianus	oproteobacteria	1 00
- — рт	NC 014707	C nsittaci	Chlamydiic	1.00	ы	NC 017140	A nastermiana-	oprotoche -t	1 00
1 11	1 110_014(9)	C. pomuci	Cinamyuna	1.00	PL E	NG_01/149	A. pusicurianus	aproteonacteria	1.00
DT	NO OFFICE	a	C111	1 1 0 0					1 0 0 -
\mathbf{PL}	NC_015217	C. psittaci	Chlamydiia	1.00	PL	NC_013859	Azospirillum sp.	α proteobacteria	0.67
$_{PL}^{PL}$	NC_015217 NC_017288	C. psittaci C. psittaci	Chlamydiia Chlamydiia	1.00 1.00	PL PL	NC_013859 NC_006672	Azospirillum sp. G. oxydans	lphaproteobacteria $lpha$ proteobacteria	0.67 0.93

PL	NC_006673	G. oxydans	α proteobacteria	1.00	PL	NC_015313	P. dioxanivorans	Actinobacteridae	0.00
PL	NC_016037	G. xylinus	α proteobacteria	1.00	\mathbf{PL}	NC_001880	A. aeolicus	Aquificae	1.00
PL	NC_014466	S. macrogoltabida	α proteobacteria	0.71	\mathbf{PL}	NC_002095	C. limicola	Chlorobia	0.83
					\mathbf{PL}	NC_010479	Synechococcus sp.	Chroobacteria	1.00
cluster	: 48 stability r	neasure : 0.969602			\mathbf{PL}	NC_010629	N. punctiforme	Homogonae	1.00
PL	NC_012552	L. johnsonii	Bacilli	1.00	\mathbf{PL}	NC_013733	S. linguale	Cytophagia	0.86
PL	NC_011998	M. caseolyticus	Bacilli	1.00	\mathbf{PL}	NC_014026	S. ruber	Cytophagia	1.00
PL	NC_012000	M. caseolyticus	Bacilli	1.00	\mathbf{PL}	NC_006907	L. ferrooxidans	Nitrospira	0.86
PL	NC_003201	O. oeni	Bacilli	1.00	\mathbf{PL}	NC_015181	A. multivorum	α proteobacteria	0.43
PL	NC_001763	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_015188	A. multivorum	α proteobacteria	1.00
PL	NC_001767	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_010123	G. diazotrophicus	α proteobacteria	1.00
PL	NC_001797	S. agalactiae	Bacilli	0.00	\mathbf{PL}	NC_010518	M. radiotolerans	α proteobacteria	1.00
PL	NC_002096	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_012634	R. africae	α proteobacteria	0.81
PL	NC_002129	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_007110	R. felis	α proteobacteria	0.00
PL	NC_005243	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_007111	R. felis	αproteobacteria	0.31
PL	NC_005564	S. aureus	Bacilli	1.00	\mathbf{PL}	NZ_CM001468	R. helvetica	α proteobacteria	0.73
PL	NC_006629	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_009897	R. massiliae	αproteobacteria	0.50
PL	NC_006977	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_016939	R. massiliae	αproteobacteria	0.59
PL	NC_007791	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_017055	R. rhipicephali	aproteobacteria	1.00
PL	NC_010111	S. aureus	Bacilli	1.00	\mathbf{PL}	NZ_CM000771	Rickettsia endosymbiont	aproteobacteria	0.31
PL	NC_010262	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_013937	C. Rickettsia	αproteobacteria	1.00
PL	NC_010284	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_013938	C. Rickettsia	aproteobacteria	0.36
PL	NC_011605	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_010982	Z. mobilis	αproteobacteria	1.00
PL	NC_013306	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_013786	Z. mobilis	aproteobacteria	1.00
PL	NC_013307	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_013784	Z. mobilis	aproteobacteria	0.43
PL	NC_013308	S. aureus	Bacilli	1.00	PL	NC_004457	Z. mobilis	oproteobacteria	1.00
PL	NC 013309	S. aureus	Bacilli	1.00	PL	NC 008763	P. nanhthalenivorans	βproteobacteria	0.95
PL	NC 013311	S aureus	Bacilli	1.00	PL	NC 013224	D rethaense	δproteobacteria	0.59
PL	NC 013312	S. aureus	Bacilli	1.00	PL	NC 002579	A.actinomucetemcomitans	γproteobacteria	0.86
PL	NC 013328	S aureus	Bacilli	1.00	PL	NC 009476	A hestiarum	vproteobacteria	0.50
PL	NC 013336	S. aureus	Bacilli	1.00	PL	NC 007100	P. aeruainosa	γproteobacteria	0.86
PL	NC 013391	S aureus	Bacilli	1.00	PL	NC 005909	P alcaligenes	vproteobacteria	1.00
PL	NC_013394	S. aureus	Bacilli	1.00	PL.	NC_011604	S. enterica	vproteobacteria	1 00
PL	NC 013452	S. aureus	Bacilli	1.00	PL.	NC 015972	S. marcescens	vproteobacteria	1 00
PL	NC 017335	S. aureus	Bacilli	1.00	PI	NC 016053	Y arboricola	oproteobacteria	0.88
PL	NC 017334	S aureus	Bacilli	1.00	PI	NC 006154	V neeudotuberculosis	oproteobacteria	0.00
PL	NC 010141	S. aureus	Bacilli	1.00	1.1	110_000104	1. pseudoraoerealosis	proteobacteria	0.50
PL	NC 005007	S. enidermidis	Bacilli	1.00	clustor	• 52 etability m	02811230		
DI DI	NC 005008	S. epidermidis	Bacilli	1.00	CUP	NC 017000	F monthematic		1.00
PL.	NC 002093	S. luadunensis	Bacilli	1.00	CHR	NC 008601	F. novicida	oproteobacteria	1.00
PL	NC 006074	S. eciuri	Bacilli	1.00	CHB	NC 017440	F of monicida	proteobacteria	1.00
PL.	NC 013033	S. simulans	Bacilli	1.00	CHR	NC 017449	F. cf. novicida	oproteobacteria	1.00
112	110_013033	D. Stitutans	Dacim	1.00	CHR	NC 010226	F. cj. noviciau	yproteobacteria	1.00
aluston	• 40 stability v	manguna + 0 855776			CHR	NC 006570	F. philomitagia	γproteobacteria	1.00
DI	149 stability I	C L		1.00	CHR	NC_000570	F. Lutarensis	γproteobacteria	1.00
PL	NC_008050	C. coli	eproteobacteria	1.00	CHR	NC-007880	F. tularensis	γ proteobacteria	1.00
PL	NC 017730	H. Cetorum	eproteobacteria	1.00	CHR	NC_008245	F. Lutarensis	γproteobacteria	1.00
	NC 008428	H. Cetorum	eproteobacteria	1.00	CHR	NC 000257	F. tularensis	γproteobacteria	1.00
	NC 006075	C. jejuni	eproteobacteria	1.00	CHR	NC 000740	F. tularensis	γproteobacteria	1.00
PL DI	NC_000975	C. lari	eproteobacteria	1.00	CHR	NC_009749	F. tularensis	γproteobacteria	1.00
PL DI	NC 008220	U. a sin snushis	eproteobacteria	1.00	CHR	NC 016022	F. tularensis	vproteobacteria	1.00
PL DI	NC_008230	H. acinonychis	eproteobacteria	1.00	CHR	NC_016933	F. tularensis	γproteobacteria	1.00
	NC 017280	H. pylori	eproteobacteria	1.00	CHR	NC 017452	F. tularensis	γproteobacteria	1.00
PL DI	NC 014161	H. pylori	eproteobacteria	1.00	CHR	NC_017455	F. tutarensis	γproteobacteria	1.00
	NC 010022	H. pylori	eproteobacteria	1.00	Спк	NC_015090	Francisella sp.	γproteobacteria	1.00
PL DI	NC 017724	H. pylori	eproteobacteria	1.00	almatan	. FO			
	NC 017064	H. pylori	eproteobacteria	1.00	Cluster	NC 010544	C h D stall	M . 11"	1.00
PL DI	NC 011224	H. pylori	eproteobacteria	0.50	DI	NC_010544	Cana. P. australiense	Mollicutes	1.00
	NC 004767	H. pylori	eproteobacteria	1.00	PL	NC_003353	P. aster yellows	Mollicutes	1.00
	NC 012547	II. pylori	epioteobacteria	1.00	DI	NC 007710	P. aster yellows	Monicutes	1.00
	NC 011400	H. pylori	eproteobacteria	1.00	PL	NC-007710	P. aster yellows	Mollicutes	nan
PL DI	NC 004040	H. pylori	eproteobacteria	1.00	PL	NC 007710	P. aster yellows	Monicutes	1.00
	NC 004949	II. pylori	epioteobacteria	1.00	PL	NC-007719	P. aster yellows	Mollicutes	1.00
PL	NC 017356	H nulori	eproteobacteria	0.64	F L DI	NC 005012	P boot	Mollieutes	1.00
рт	NC 014957	H nulori	eproteobacteria	1.00	рт Т	NC 016F92	P Brassica	Mollieutes	1.00
PL	NC 0014257	H nulori	eproteobacteria	1.00	CHB	NC 005202	P opiop vollows	Mollieutes	1.00
PL	NC 004845	H nulori	eproteobacteria	1.00	DI	NC 006002	P onion vellews	Mollicutes	1.00
рт	NC 010894	H. pytori	eproteobacteria	0.73	гъ рт	NC 012080	P onion vellews	Mollieutes	1.00
	NC 002110	H. pylori	eproteobacteria	1.00		NC_012089	P. onion yellows	Malliantes	1.00
рт	NC 017979	H. pytori	eproteobacteria	1.00	гъ рт	NC 010160	P onion vellews	Mollieutes	1.00
DI	NC 017970	H pulori	cproteobacteria	0.50	г L DT	NC 010160	P opiop volle	Mollioutes	1.00
PL DI	NC 017277	H palori	eproteobacteria	0.50		NC 010170	P. onion yellows	Malliantes	1.00
рт	NC 004007	n. pyton C. jejuni	eproteobacteria	1.00	гь рт	NC 010171	P onion vellews	Mollicutes	1.00
PL DI	NC 017260	C. jejuni H. pulori	eproteobacteria	1.00		NC 010405	P. Davisonia	Malliantes	1.00
PI.	NC 017364	H pylori	eproteobacteria	1.00	гь рт	NC 010405	P Pauloursia	Mollieutes	1.00
DI	NC 017262	H palori	cproteobacteria	1.00		NC 004800	P. Paulowilla	Malliantes	1.00
PT.	NC 008087	H pylori	eproteobacteria	1.00	PL DI	NC 014122	P Behmannia	Mollicutes	1.00
рт	NC 001756	H nulori	eproteobacteria	1.00	рт I	NC 010020	P tomate	Mollieutes	1.00
PT.	NC 010113	Vibrio sp	oproteobacteria	0.67	гL	110-010920	1. tomato	monicutes	0.80
·		, 10110 sp.	/Proteobacteria	0.01	alustar	. 54 stability	assura · 1 000000		
cluster	: 50 stability r	neasure : 0.675132			PI.	NC 011340	B cereus	Bacilli	1.00
CHR	NC 012526	D deserti	Deinococci	1.00	PT.	NC 018404	B cereus	Bacilli	1.00
CHR	NC 008025	D. geothermalis	Deinococci	1.00	PI.	NC 001004	S aureus	Bacilli	1 00
PL	NC 000020	D geothermalie	Deinococci	0.36	DI	NC 001005	S. aureus	Bacilli	1.00
PI.	NC 017771	D aphiensis	Deinococci	0.57	PT.	NC 005565	S aureus	Bacilli	1.00
СНВ	NC 017700	D. gobiensis	Deinococci	0.67	DI	NC 007200	S. aureus	Bacilli	1.00
pr	NC 017701	D. goviensis	Deinococci	0.57	гъ рт	NC 010695	S. aureus	Bacilli	1.00
Снь	NC 014059	D. youchists	Deinococci	1.00	гь рт	NC 010696	S. aureus	Bacilli	1.00
CHP	NC 015161	D. mancopensis	Deinococci	1.00	гъ рт	NC 01220F	S. aureus	Bacilli	1.00
DI	NC 015162	D. proteolyticus	Deinococci	0.53	г L DT	NC 012210	S. aureus	Dacilli	1.00
Снь	NC 001962	D. proteotyticus	Deinococci	0.00	гъ рт	NC 012201	S. aureus	Bacilli	1.00
	110_001203	D. Tuutouutuus	Demototti	1 11411	PL PL	NC 013276	S. aureus	Bacilli	1.00
aluate	• 51 ctob:1:4-	nossuro · 0 600020			г L рт	NC 018060	S. aureus	Dacill:	1.00
ciuster	• •• stability f	neasure . 0.000002			гL	110-010909	o. aureus	Dacini	1 1.00

	NC_019139	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_017806	D. gobiensis	Deinococci	1.00
\mathbf{PL}	NC_019143	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_008010	D. geothermalis	Deinococci	1.00
\mathbf{PL}	NC_019145	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_015163	D. proteolyticus	Deinococci	0.88
\mathbf{PL}	NC_019147	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_015169	D. proteolyticus	Deinococci	1.00
\mathbf{PL}	NC_007170	S. haemolyticus	Bacilli	1.00	\mathbf{PL}	NC_015170	D. proteolyticus	Deinococci	1.00
\mathbf{PL}	NC_006871	S. lentus	Bacilli	1.00	\mathbf{PL}	NC_000958	D. radiodurans	Deinococci	1.00
\mathbf{PL}	NC_002059	B. fibrisolvens	Clostridia	1.00	RECE	NC_001264	D. radiodurans	Deinococci	1.00
PL	NC_004977	$S. \ ruminantium$	Negativicutes	1.00	PL	NC_014213	M. silvanus	Deinococci	1.00
PL	NC_006857	S. ruminantium	Negativicutes	1.00	PL	NC_014214	M. silvanus	Deinococci	1.00
PL	NC_016045	S. ruminantium	Negativicutes	1.00	PL	NC_014753	O. profundus	Deinococci	1.00
PL	NC_008049	C. coli	eproteobacteria	1.00	PL	NC_019387	T. oshimai	Deinococci	0.97
PL	NC_008153	uncultured	Human gut	1.00	PL	NC_019388	T. oshimai	Deinococci	0.95
alwatan .	EE				PL	NC-005838	1. thermophilus	Deinococci	1.00
CUD	NC 007760	A d-b-d	Stbti-	1.00	F L DI	NC 017500	T. thermophilus	Deinococci	1.00
CHR	NC 011801	A. debalogenans	aproteobacteria	1.00	PL	NC 017273	T thermonhilus	Deinococci	0.95
CHR	NC 009675	Anaeromyrobacter sp	δproteobacteria	1.00		1101011210	11 00010000000	Democoder	0.000
CHR	NC 011145	Anaeromyzobacter sp.	δproteobacteria	1.00	cluster	60 stability m	neasure : 0.451109		
				1	PL	NC_010474	Synechococcus sp.	Chroobacteria	0.23
cluster :	56 stability m	neasure : 0.979487			PL	NC_005230	Synechocystis sp.	Chroobacteria	0.14
PL	NC_015509	B. cecembensis	Bacilli	1.00	\mathbf{PL}	NC_015179	A. multivorum	α proteobacteria	0.92
\mathbf{PL}	NC_015849	E. faecium	Bacilli	1.00	\mathbf{PL}	NC_001275	A. aceti	α proteobacteria	0.94
\mathbf{PL}	NC_017023	E. faecium	Bacilli	1.00	\mathbf{PL}	NC_009470	A. cryptum	α proteobacteria	0.64
\mathbf{PL}	NC_001272	B. thuringiensis	Bacilli	1.00	PL	NC_015462	Cand. Rickettsia	α proteobacteria	0.33
\mathbf{PL}	NC_002091	B. thuringiensis	Bacilli	1.00	\mathbf{PL}	NC_010514	M. radiotolerans	α proteobacteria	0.64
\mathbf{PL}	NC_004335	B. thuringiensis	Bacilli	1.00	PL	NC_007641	R. rubrum	α proteobacteria	0.77
PL	NC_011796	B. thuringiensis	Bacilli	1.00	PL	NC_006675	G. oxydans	α proteobacteria	0.25
PL	NC_017207	B. thuringiensis	Bacilli	1.00	PL	NC_019397	G. oxydans	α proteobacteria	0.88
PL	NC_017209	B. thuringiensis	Bacilli	1.00	PL	NC_016028	G. xylinus	αproteobacteria	nan
PL	NC_017211	B. thuringiensis	Bacilli	1.00	PL	NC_012989	M. extorquens	αproteobacteria	0.30
	NC_018882	B. thuringiensis	Bacilli	1.00		NC 017041	M. noaulans		0.12
	NC 005242	B. thuringiensis	Bacilli	1.00	PL PI	NC 010027	R. austratis R. monacensis	aproteobacteria	0.75
PL	NC 012003	L. sphuericus M. caseoluticus	Bacilli	1.00	PL	NC 005297	R palustris	aproteobacteria	0.37
PL.	NC 002094	M. caseolyncus M. halophilus	Bacilli	1.00	PL	NC 015715	Z mohilis	aproteobacteria	0.12
PL	NC 002013	S. aureus	Bacilli	1.00	PL	NC_013190	C. Accumulibacter	βproteobacteria	0.43
PL	NC_010426	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_008341	N. eutropha	β proteobacteria	0.12
\mathbf{PL}	NC_010427	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_008342	N. eutropha	β proteobacteria	0.94
\mathbf{PL}	NC_010616	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_014154	T.~intermedia	β proteobacteria	0.59
\mathbf{PL}	NC_013314	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_014755	S. kujiense	ϵ proteobacteria	0.88
\mathbf{PL}	NC_013346	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_014756	S. kujiense	ϵ proteobacteria	0.71
\mathbf{PL}	NC_019011	S. aureus	Bacilli	1.00	\mathbf{PL}	NC_005023	A. ferrooxidans	γ proteobacteria	0.57
\mathbf{PL}	NC_019140	S. aureus	Bacilli	1.00	PL	NC_012961	P. asymbiotica	γ proteobacteria	0.64
PL	NC_007768	S. chromogenes	Bacilli	1.00	PL	NC_010891	Pseudomonas sp.	γ proteobacteria	0.92
PL	NC_008354	S. chromogenes	Bacilli	1.00	PL	NC_019113	S. enterica	γ proteobacteria	0.77
PL	NC_003969	S. epidermidis	Bacilli	1.00	PL	NC_017557	X. albilineans	γ proteobacteria	0.50
PL	NC_007351	S. saprophyticus	Bacilli	1.00	almatan	61			
PL	NC_010626	S. sciuri	Bacilli	1.00	Cluster	61 stability ff	leasure : 0.027095		
DI	I NC 015176	S aimulana	Regula	1 1 00		NC 000026	A manina	Chrochasteria	0.05
PL PL	NC_015176	S. simulans S. thermorphilus	Bacilli	1.00	PL	NC_009926	A. marina	Chroobacteria Chroobacteria	0.95
PL PL PL	NC_015176 NC_004968 NC 005208	S. simulans S. thermophilus S. warneri	Bacilli Bacilli Bacilli	1.00 1.00 1.00	PL PL PL	NC_009926 NC_009928 NC_009929	A. marina A. marina A. marina	Chroobacteria Chroobacteria Chroobacteria	0.95 0.97 1.00
PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165	S. simulans S. thermophilus S. warneri S. warneri	Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00	PL PL PL PL	NC_009926 NC_009928 NC_009929 NC_009930	A. marina A. marina A. marina A. marina	Chroobacteria Chroobacteria Chroobacteria Chroobacteria	0.95 0.97 1.00 1.00
PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166	S. simulans S. thermophilus S. warneri S. warneri S. warneri	Bacilli Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL	NC_009926 NC_009928 NC_009929 NC_009930 NC_009933	A. marina A. marina A. marina A. marina A. marina	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	$\begin{array}{c} 0.95 \\ 0.97 \\ 1.00 \\ 1.00 \\ 0.93 \end{array}$
PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073	S. simulans S. thermophilus S. warneri S. warneri B. fragilis	Bacilli Bacilli Bacilli Bacilli Bacilli Bactli	$ \begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array} $	PL PL PL PL PL PL	NC_009926 NC_009928 NC_009929 NC_009930 NC_009933 NC_011885	A. marina A. marina A. marina A. marina A. marina Cyanothece sp.	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	$\begin{array}{c} 0.95 \\ 0.97 \\ 1.00 \\ 1.00 \\ 0.93 \\ 0.96 \end{array}$
PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073	S. simulans S. thermophilus S. warneri S. warneri B. fragilis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00 \end{array} $	PL PL PL PL PL PL CHR	NC_009926 NC_009928 NC_009929 NC_009930 NC_009933 NC_011885 NC_009052	A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria	$\begin{array}{c} 0.95 \\ 0.97 \\ 1.00 \\ 1.00 \\ 0.93 \\ 0.96 \\ 0.79 \end{array}$
PL PL PL PL PL PL cluster :	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 : 57 stability m	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00 \end{array} $	PL PL PL PL PL CHR CHR	NC_009926 NC_009928 NC_009929 NC_009930 NC_009933 NC_011885 NC_009052 NC_017579	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.95 \\ 0.97 \\ 1.00 \\ 1.00 \\ 0.93 \\ 0.96 \\ 0.79 \\ 0.21 \end{array}$
PL PL PL PL PL PL cluster : PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability m NC_010905	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica	Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 0.69	PL PL PL PL PL CHR CHR PL	NC_009926 NC_009928 NC_009930 NC_009933 NC_011885 NC_009052 NC_017579 NC_017580	A. marina A. marina A. marina A. marina A. marina S. baltica S. baltica S. baltica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.95 \\ 0.97 \\ 1.00 \\ 1.00 \\ 0.93 \\ 0.96 \\ 0.79 \\ 0.21 \\ 0.63 \end{array}$
PL PL PL PL PL PL Cluster : PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability m NC_010905 NC_018580	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp.	Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54	PL PL PL PL PL CHR CHR PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.011885 NC.010852 NC.017579 NC.017580	A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.95 \\ 0.97 \\ 1.00 \\ 1.00 \\ 0.93 \\ 0.96 \\ 0.79 \\ 0.21 \\ 0.63 \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 : 57 stability m NC_010905 NC_018580 NC_009660	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans	Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.69\\ 0.54\\ 0.54\\ 0.54\\ 0.54 \end{array} $	PL PL PL PL PL CHR CHR PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.009933 NC.011885 NC.010952 NC.017579 NC.017580	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.95 \\ 0.97 \\ 1.00 \\ 1.00 \\ 0.93 \\ 0.96 \\ 0.79 \\ 0.21 \\ 0.63 \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability m NC_018580 NC_009660 NC_018022 NC_018022	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.69\\ 0.54\\ 0.54\\ 0.54\\ 0.70\\ 0.86\end{array}$	PL PL PL PL PL PL CHR PL CHR PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012555	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria proteobacteria proteobacteria proteobacteria	$\begin{array}{c} 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.96\\ 0.79\\ 0.21\\ 0.63\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 : 57 stability m NC_010905 NC_0180580 NC_009660 NC_018022 NC_008623	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense	Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.86\\ 1.00\\ 0.86\\ 0.96$	PL PL PL PL PL CHR CHR PL CHR PL PL PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.00229	A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. cloacae	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability m NC_010905 NC_018580 NC_008600 NC_018022 NC_018023 NC_008562	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense	Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.69\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ \end{array}$	PL PL PL PL PL CHR CHR PL CHR PL PL PL PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012556 NC.009838 NC 013365	A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. coli E. coli	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.96\\ 0.79\\ 0.21\\ 0.63\\ \hline \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability n NC_010905 NC_018580 NC_00860 NC_018022 NC_018023 NC_008703 NC_006362 NC_005314	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense Mycobacterium sp. N. farcinica P. dioraniyorans	Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ \end{array}$	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012556 NC.012556 NC.009838 NC.013365 NC.013365	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli E. coli E. coli	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.96\\ 0.79\\ 0.21\\ 0.63\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability n NC_010905 NC_018580 NC_008600 NC_018022 NC_018023 NC_008703 NC_006362 NC_005314 NC_002576	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi	Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ 0.57\\ \end{array}$	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017580 62 stability m NC.012555 NC.012556 NC.012556 NC.010838 NC.01365 NC.016980	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli K. pneumoniae K. pneumoniae	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.96\\ 0.71\\ 0.63\\ \hline \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_011073 57 stability n NC_010905 NC_018580 NC_009660 NC_018022 NC_018023 NC_006362 NC_006362 NC_004854	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense R. equi P. dioxanivorans R. equi	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ 0.57\\ 0.38\\ \end{array}$	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012556 NC.00838 NC.013365 NC.010870 NC.010870 NC.010870 NC.010880	A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli K. pneumoniae K. pneumoniae S. enterica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.95 0.97 1.00 1.00 0.93 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_01165 NC_011073 57 stability m NC_010905 NC_018022 NC_018022 NC_008600 NC_018023 NC_006362 NC_016314 NC_002576 NC_004854 NC_011150	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ 0.57\\ 0.38\\ 0.64\\ \end{array}$	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.013365 NC.013365 NC.013365 NC.013305 NC.016980 NC.003305 NC.003384	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability m NC_010905 NC_01905 NC_009660 NC_018022 NC_018022 NC_018023 NC_006362 NC_006362 NC_004854 NC_001151	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.69\\ 0.54\\ 0.54\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ 0.57\\ 0.38\\ 0.64\\ 0.64\\ \end{array}$	PL PL PL PL PL CHR CHR PL CHR PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012555 NC.012556 NC.009838 NC.010870 NC.016980 NC.002305 NC.003384 NC.009981	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability m NC_010905 NC_018020 NC_018022 NC_018022 NC_018023 NC_006362 NC_004854 NC_004854 NC_004854 NC_0011511 NC_007486	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense R. chubuense R. equi R. equi R. equi R. equi R. equi R. erythropolis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.69\\ 0.54\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ 0.57\\ 0.38\\ 0.64\\ 0.64\\ 0.64\\ 0.64\\ \end{array}$	PL PL PL PL PL CHR CHR PL CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.000952 NC.017579 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012556 NC.008388 NC.010870 NC.016880 NC.003384 NC.009981 NC.016825	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli E. coli E. coli E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica S. enterica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007166 NC_011073 57 stability m NC_010905 NC_018023 NC_018023 NC_008703 NC_008703 NC_006362 NC_015314 NC_002576 NC_004854 NC_011150 NC_007486 NC_012523	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equis R. opacus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.017580 62 stability m NC.012555 NC.012556 NC.012556 NC.0108388 NC.010870 NC.016980 NC.003384 NC.003384 NC.00384 NC.00384 NC.003825 NC.016825 NC.019114	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica Deasure : 0.919630 E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica S. enterica S. enterica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.96\\ 0.79\\ 0.21\\ 0.63\\ \hline \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_011073 57 stability m NC_010905 NC_018580 NC_009660 NC_018022 NC_018022 NC_008703 NC_006362 NC_004854 NC_004854 NC_011150 NC_011151 NC_007486 NC_012523 NC_013420	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense M. chubuense R. caliotolerans R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equis Streptomyces sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.69\\ 0.54\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ 0.57\\ 0.38\\ 0.64\\ 0.64\\ 1.00\\ 0.82\\ \end{array}$	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.017579 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.013365 NC.00838 NC.013365 NC.003384 NC.003384 NC.009810 NC.009810 NC.003384 NC.009811 NC.019114 NC.004989	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	0.95 0.97 1.00 1.00 0.93 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_011073 : 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_008600 NC_018022 NC_006362 NC_006362 NC_006362 NC_004854 NC_0011551 NC_007486 NC_012523 NC_013420 NC_013420 NC_014159	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi Streptomyces sp. T. paurometabola	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.58 0.57 0.88 0.57 0.68 0.64 0.64 1.00 0.82 nan	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012555 NC.013365 NC.013365 NC.013365 NC.013365 NC.013870 NC.016980 NC.003384 NC.00384 NC.00384 NC.004989 NC.004989 NC.005211	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens S. marcescens	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.95\\ 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.93\\ 0.79\\ 0.21\\ 0.63\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC.015176 NC.004968 NC.005208 NC.005208 NC.007165 NC.011073 57 stability m NC.010905 NC.01905 NC.018022 NC.018022 NC.018022 NC.018022 NC.018023 NC.006362 NC.004354 NC.004854 NC.011551 NC.011551 NC.014159 NC.014028 NC.014028	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chu	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012555 NC.01365 NC.013865 NC.010870 NC.016980 NC.003384 NC.009981 NC.016825 NC.019114 NC.004989 NC.005211	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. marcescens S. marcescens	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria Pproteobacteria	$\begin{array}{c} 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.79\\ 0.21\\ 0.63\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007165 NC_011073 57 stability m NC_010905 NC_01905 NC_00860 NC_018022 NC_018022 NC_018023 NC_006362 NC_004854 NC_0015314 NC_002576 NC_004854 NC_011151 NC_007486 NC_012523 NC_014159 NC_014159 NC_014028 NC_01372	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense Streptomyces sp. T. paurometabola S. ruber M. fulvus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.70 0.86 1.00 0.90 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.18 nan 0.18 nan 0.20 0.20 0.10 0.10 0.10 0.10 0.10 0.54 0.57 0.38 0.64 0.64 0.52 0.58 0.54 0.54 0.57 0.38 0.54 0.54 0.54 0.57 0.38 0.54 0.54 0.54 0.57 0.38 0.54 0.54 0.54 0.57 0.38 0.54 0.54 0.54 0.57 0.38 0.64 0.58 0.58 0.57 0.58 0.54 0.54 0.54 0.57 0.38 0.54 0.54 0.54 0.54 0.57 0.38 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.58 0.5	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.000952 NC.017579 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012556 NC.009838 NC.010870 NC.016870 NC.00384 NC.00981 NC.016825 NC.019114 NC.004989 NC.005211 : 63 stability m	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli E. coli E. coli E. coli E. coli E. coli E. coli E. coli S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens S. marcescens measure : 0.668620	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria pproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007165 NC_011073 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_018022 NC_018023 NC_006362 NC_004854 NC_001314 NC_002576 NC_004854 NC_011311 NC_007486 NC_011311 NC_007486 NC_012523 NC_013129 NC_014028 NC_014028 NC_010372 NC_013775	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi S. reptorpolis S. ruber M. fulvus P. damselae	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.70 0.86 1.00 0.90 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017580 : 62 stability m NC.012555 NC.012556 NC.008388 NC.013365 NC.010870 NC.00384 NC.00384 NC.00981 NC.016825 NC.019114 NC.004989 NC.005211 : 63 stability m NC.009012 NC.009012	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica Measure : 0.919630 E. cloacae E. cloacae E. cloacae E. coli E. coli E. coli E. coli E. coli E. coli E. coli E. coli S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens marcescens marcescens Measure : 0.668620 C. thermocellum	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	0.95 0.97 1.00 1.00 0.93 0.93 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_01073 : 57 stability m NC_010905 NC_018580 NC_009660 NC_018022 NC_018023 NC_006362 NC_008703 NC_006362 NC_01314 NC_002576 NC_004854 NC_011150 NC_01150 NC_01151 NC_007486 NC_012523 NC_013420 NC_014159 NC_014028 NC_010372 NC_013775	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense S. raterina R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi Streptomyces sp. T. paurometabola S. ruber M. fulvus P. damselae measure : 0.420319	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.69\\ 0.54\\ 0.54\\ 0.70\\ 0.86\\ 1.00\\ 0.90\\ 0.68\\ 0.57\\ 0.38\\ 0.64\\ 0.64\\ 1.00\\ 0.82\\ nan\\ 0.18\\ nan\\ 0.69\\ \end{array}$	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.013365 NC.00838 NC.013365 NC.00838 NC.013365 NC.003384 NC.002305 NC.003384 NC.009810 NC.009811 NC.009811 NC.009811 NC.009811 NC.009811 NC.009811 NC.009114 NC.009012 NC.017304	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica Measure : 0.919630 E. cloacae E. cloacae E. cloacae E. coli E. coli K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens S. marcescens marcescens Measure : 0.668620 C. thermocellum C. thermocellum	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.95\\ 0.95\\ 0.97\\ 1.00\\ 1.00\\ 0.93\\ 0.93\\ 0.79\\ 0.21\\ 0.63\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_011073 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_018022 NC_00860 NC_008023 NC_006362 NC_004854 NC_011150 NC_001459 NC_01323 NC_013420 NC_013420 NC_01372 NC_01372 S8 stability m NC 000620	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. opacus Streptomyces sp. T. paurometabola S. ruber M. fulvus P. damselae measure : 0.420319	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.70 0.86 1.00 0.90 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.82 nan 0.82 nan 0.69 0.54 0.69 0.54 0.64 0.64 0.64 0.64 0.64 0.64 0.69 0.82 nan 0.82 nan 0.82 nan 0.69 0.54 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.64 0.69 0.82 0.82 0.82 0.69 0.69 0.64 0.64 0.64 0.64 0.64 0.64 0.69 0.82 0.82 0.82 0.82 0.82 0.82 0.69 0.69 0.68 0.64 0.64 0.64 0.69 0.82 0.82 0.82 0.86 0.82 0.82 0.82 0.82 0.82 0.82 0.82 0.82 0.82 0.82 0.82 0.84 0.69 0.82 0.82 0.82 0.84 0.69 0.82 0.82 0.82 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.69 0.69 0.69 0.69 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.69 0.69 0.69 0.82 0.82 0.69	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012555 NC.013365 NC.013365 NC.013365 NC.013365 NC.013365 NC.013365 NC.013384 NC.016980 NC.003384 NC.00384 NC.003981 NC.016825 NC.019114 NC.004989 NC.005211 : 63 stability m	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae S. enterica S. marcescens S. marcescens E. clo68620 C. thermocellum C. thermocellum	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007165 NC_011073 : 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_008362 NC_006362 NC_004854 NC_011551 NC_004854 NC_01151 NC_004854 NC_01151 NC_014159 NC_013420 NC_013420 NC_01375 : 58 stability m NC_003078	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi S. truber M. fulvus P. damselae measure : 0.420319 S. medicae S. mediloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.56 0.58 0.68 0.57 0.882 nan 0.18 nan 0.69 1.00 0.69 0.69 0.69 0.54 0.54 0.54 0.54 0.54 0.57 0.82 0.59 0.10 0.82 0.59 0.10 0.69 0.54 0.54 0.57 0.82 0.59 0.18 0.69 0.69 0.69 0.69 0.69 0.69 0.69 0.64 0.64 0.69 0.69 0.69 0.69 0.69 0.68 0.57 0.82 0.69 0.69 0.69 0.69 0.69 0.64 0.69 0.69 0.69 0.69 0.64 0.69 0.60 0.60 0.69 0.60 0.69 0.60 0.60 0.60 0.69 0.60 0.60 0.60 0.60 0.69 0.60 0	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012555 NC.013365 NC.013365 NC.013365 NC.013365 NC.013365 NC.013384 NC.016980 NC.003384 NC.003384 NC.003981 NC.016825 NC.003384 NC.004989 NC.005211 : 63 stability m NC.009012 NC.017304 : 64 stability m	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica Cyanothece sp. S. baltica S. baltica E. colo E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. marcescens S. marcescens S. marcescens Marcescens	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC.015176 NC.004968 NC.005208 NC.005208 NC.007165 NC.011073 : 57 stability m NC.010905 NC.01905 NC.018022 NC.018022 NC.018022 NC.018022 NC.018022 NC.018023 NC.006362 NC.004854 NC.004854 NC.011151 NC.007486 NC.012523 NC.014159 NC.014159 NC.014028 NC.013775 : 58 stability m NC.008620 NC.00378 NC.00378	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi R. equi Streptomyces sp. T. paurometabola S. ruber M. fulvus P. damselae measure : 0.420319 S. medicae S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteriae Actinobacteriae Actinobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.68 0.57 0.38 0.64 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69 0.58 0.58 0.54 0.64 0.64 0.69 0.58 0.54 0.64 0.64 0.69 0.58 0.58 0.54 0.64 0.64 0.69 0.18 nan 0.69 0.69 0.18 nan 0.69 0.18 0.18 nan 0.69 0.18 0.10 0.18 0.18 0.10 0.00 0.18 0.18 0.10 0.18 0.10 0.18 0.18 0.10 0.10 0.18 0.10 0.18 0.10 0.18 0.10 0.10 0.18 0.10 0.18 0.10 0.10 0.18 0.10 0.10 0.18 0.10 0.10 0.18 0.10 0.10 0.18 0.10 0.10 0.18 0.10 0.10 0.18 0.10 0.10 0.18 0.100 0.100 0.18 0.100 0.100 0.18 0.100 0.100 0.100 0.18 0.100 0.100 0.100 0.100 0.18 0.100 0.0000 0.000 0.000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000000	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009930 NC.011885 NC.000952 NC.017579 NC.017579 NC.017580 : 62 stability n NC.012555 NC.012556 NC.009838 NC.010870 NC.010870 NC.003384 NC.009981 NC.016825 NC.019114 NC.009981 NC.016825 NC.019114 NC.009981 NC.004829 NC.005211 : 63 stability n NC.007304 : 64 stability n NC.003983	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens S. marcescens S. marcescens Marcescens Marcescelum C. thermocellum C. thermocellum D. shibae H. baltica	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_011073 57 stability m NC_010905 NC_018580 NC_008600 NC_018022 NC_018023 NC_008703 NC_006362 NC_004854 NC_01150 NC_011150 NC_011150 NC_01151 NC_004854 NC_011523 NC_013420 NC_013420 NC_01372 NC_013775 58 stability m NC_009620 NC_003078 NC_003078 NC_003078 NC_003078	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi R. equi R. equi S. reptomyces sp. T. paurometabola S. ruber M. fulvus P. damselae measure : 0.420319 S. medicae S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.54 0.70 0.86 1.00 0.68 0.57 0.38 0.64 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69 1.00	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.017579 NC.012555 NC.012556 NC.009838 NC.008370 NC.016805 NC.009881 NC.00981 NC.00981 NC.00981 NC.00981 NC.00981 NC.00981 NC.00981 NC.009211 : 63 stability m NC.009012 NC.017304 : 64 stability m NC.00956 NC.007801	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica Measure : 0.919630 E. cloacae E. cloacae E. coli E. coli K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens Marcescens S. marcescens Marcescens	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_011073 : 57 stability m NC_010905 NC_018022 NC_018022 NC_008600 NC_018022 NC_008703 NC_006362 NC_004854 NC_01150 NC_01150 NC_01150 NC_01423 NC_013420 NC_013420 NC_01372 S8 stability m NC_009620 NC_003078 NC_017326	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi R. equi R. equi S. treptomyces sp. T. paurometabola S. ruber M. fulvus P. damselae measure : 0.420319 S. medicae S. meliloti S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.70 0.86 1.00 0.90 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69 1.00	PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012556 NC.009838 NC.013365 NC.013365 NC.013365 NC.013365 NC.003384 NC.009980 NC.003384 NC.009981 NC.009012 NC.017304 : 64 stability m NC.009956 NC.012983 NC.012983 NC.012983 NC.012983 NC.012983	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. marcescens marcescens marcescens E. coli C. thermocellum C. thermocellum C. thermocellum C. thermocellum D. shibae H. baltica Jannaschia sp. P. aminophilus	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007165 NC_011073 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_018022 NC_008362 NC_006362 NC_006362 NC_006362 NC_006362 NC_004854 NC_011150 NC_004854 NC_011151 NC_004854 NC_011150 NC_014159 NC_013420 NC_013420 NC_013420 NC_01375 58 stability m NC_003078 NC_017323 NC_017326 NC_017326 NC_018701	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. meliloti S. meliloti S. meliloti S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria aproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.55 0.38 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.82 nan 0.18 nan 0.69 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.82 nan 0.18 nan 0.69 1.00	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012555 NC.013365 NC.013365 NC.013365 NC.013365 NC.013384 NC.003384 NC.016825 NC.019114 NC.004989 NC.005211 : 63 stability m NC.009912 NC.017304 : 64 stability m NC.009956 NC.012983 NC.007801 NC.013513 NC.018288	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. cloi K. pneumoniae K. pneumoniae S. enterica S. marcescens marcescens E. clo68620 C. thermocellum C. thermocellum C. thermocellum D. shibae H. baltica Jannaschia sp. P. aminophilus P. gallacciensis	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007165 NC_011073 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_018023 NC_006362 NC_004854 NC_011551 NC_014159 NC_014159 NC_014159 NC_01372 NC_013775 58 stability m NC_003078 NC_017323 NC_017326 NC_017321 NC_017321 NC_017321	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi S. treberomyces sp. T. paurometabola S. ruber M. fulvus P. damselae mediloti S. meliloti S. meliloti S. meliloti S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69 1.00	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012555 NC.013365 NC.013365 NC.013365 NC.013365 NC.013384 NC.009981 NC.016825 NC.003384 NC.009981 NC.016825 NC.003844 NC.009981 NC.004989 NC.005211 63 stability m NC.009012 NC.017304 64 stability m NC.009956 NC.012983 NC.012983 NC.018422	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae K. pneumoniae S. enterica S. marcescens Marcescens Marcescens Marcescens Marcescens Marcescens Marcescens Marcescens Marcescens Marcescens D. shibae H. baltica Jannaschia sp. P. gallacciensis P. gallacciensis	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC.015176 NC.004968 NC.005208 NC.005208 NC.007165 NC.011073 57 stability m NC.010905 NC.018022 NC.018022 NC.018022 NC.018022 NC.018022 NC.018022 NC.018023 NC.006362 NC.004854 NC.004854 NC.011151 NC.004854 NC.011151 NC.007486 NC.012523 NC.01372 NC.014159 NC.014028 NC.014159 NC.014028 NC.013775 58 stability m NC.008620 NC.015596 NC.017323 NC.017326 NC.01731	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. capui R. equi R. equi R. equi R. equi R. equi R. equi R. equi S. maurometabola S. meliloti S. meliloti S. meliloti S. meliloti measure : 0.756468	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria αproteobacteria αproteobacteria αproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.54 0.70 0.86 1.00 0.68 0.57 0.38 0.64 0.64 0.64 1.00 0.82 man 0.18 man 0.69 1.00	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009930 NC.011885 NC.000952 NC.017579 NC.017579 NC.017580 : 62 stability n NC.012555 NC.012556 NC.012556 NC.009838 NC.010870 NC.003305 NC.003384 NC.009981 NC.016825 NC.019114 NC.009981 NC.016825 NC.019114 NC.009981 NC.005211 : 63 stability n NC.005211 : 63 stability n NC.007304 : 64 stability n NC.012983 NC.007801 NC.018288 NC.018422 NC.018422 NC.009753	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens S. marcescens S. marcescens S. marcescens Marcescens Marcescens D. shibae H. baltica Jannaschia sp. P. aminophilus P. gallaeciensis P. gallaeciensis P. methylutens	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_01073 : 57 stability m NC_010905 NC_018580 NC_008600 NC_018022 NC_018023 NC_006362 NC_008703 NC_006362 NC_01314 NC_002576 NC_004854 NC_011150 NC_01150 NC_01151 NC_004854 NC_011523 NC_01420 NC_014159 NC_014028 NC_01372 NC_013775 : 58 stability m NC_009620 NC_003078 NC_017326 NC_017326 NC_017326 NC_019428	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense N. farcinica P. dioxanivorans R. equi R. equi R. equi R. equi R. equi R. equi R. equi S. reptomyces sp. T. paurometabola S. ruber M. fulvus P. damselae measure : 0.420319 S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.68 0.57 0.38 0.64 0.64 0.64 1.00 0.82 nan 0.69 1.00 1	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.013365 NC.00838 NC.013365 NC.013365 NC.003384 NC.009810 NC.009810 NC.009012 NC.017304 64 stability m NC.009956 NC.012983 NC.013513 NC.018288 NC.018288 NC.018288 NC.018288 NC.0182729	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica Measure : 0.919630 E. cloacae E. cloacae E. cloi E. coli E. coli E. coli K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens marcescens S. marcescens Measure : 0.668620 C. thermocellum C. thermocellum C. thermocellum C. thermocellum Measure : 0.706531 D. shibae H. baltica Jannaschia sp. P. gallaeciensis P. gallaeciensis P. gallaeciensis P. methylutens R. litoralis	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL PL PL PL PL PL PL PL PL	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_011073 57 stability m NC_010905 NC_018022 NC_018022 NC_018023 NC_006362 NC_018023 NC_006362 NC_01314 NC_002576 NC_004854 NC_011150 NC_011150 NC_01423 NC_014596 NC_01372 NC_013775 58 stability m NC_017326 NC_017326 NC_017326 NC_019428 NC_019428 NC_003267	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi S. treber M. fulvus P. damselae measure : 0.420319 S. medicae S. meliloti S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.54\\ 0.54\\ 0.54\\ 0.54\\ 0.54\\ 0.54\\ 0.64\\ 0.64\\ 0.64\\ 0.64\\ 1.00\\ 0.82\\ nan\\ 0.68\\ 0.64\\ 1.00\\ 0.82\\ nan\\ 0.69\\ \hline \end{array}$	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012555 NC.013365 NC.013365 NC.013365 NC.013870 NC.016980 NC.003384 NC.016825 NC.013842 NC.004989 NC.004989 NC.004989 NC.004989 NC.004981 NC.004989 NC.004981 NC.004983 NC.015211 : 64 stability m NC.009012 NC.017304 : 64 stability m NC.00956 NC.013513 NC.018288 NC.018288 NC.018288 NC.018729 NC.007490	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae S. enterica S. marcescens marcescens Marcescens D. shibae H. baltica Jannaschia sp. P. aminophilus P. gallaeciensis P. gallaeciensis P. methylutens R. litoralis R. sphaeroides	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL	NC_015176 NC_004968 NC_005208 NC_007165 NC_007165 NC_011073 : 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_018023 NC_006362 NC_004854 NC_004854 NC_011150 NC_014159 NC_014159 NC_013420 NC_013420 NC_013420 NC_013420 NC_01375 : 58 stability m NC_003078 NC_017326 NC_01326NC_01326 NC_01326 NC_01326NC_01326 NC_01326NC_01326 NC_01326 NC_01326NC_01326 NC_01326 NC_01326NC_01326 NC_0136NC_0136 NC_0136NC_0136 NC_0136NC_0136 NC_0136 NC_0136N	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi S. treberomyces sp. T. paurometabola S. ruber M. fulvus P. damselae measure : 0.420319 S. mediloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti Mabaena sp. Nostoc sp. Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria Aproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.50 0.90 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69 1.00 1.00 1.00 1.00 1.00 1.00 0.82 nan 0.18 nan 0.69 0.59 0.51 0.70 0.82 nan 0.18 nan 0.69 0.59 0.70 0.82 nan 0.18 nan 0.69 0.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.82 nan 0.18 nan 0.69 0.59 0.71 0.00 1.00	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012556 NC.009838 NC.013365 NC.013365 NC.013365 NC.013365 NC.013848 NC.003844 NC.009981 NC.016825 NC.013844 NC.004989 NC.005211 63 stability m NC.009912 NC.017304 64 stability m NC.009956 NC.012983 NC.007801 NC.018288 NC.018288 NC.018422 NC.007523 NC.007490 NC.007489	A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. cloi E. coli K. pneumoniae S. enterica S. marcescens marcescens D. shibae H. baltica Jannaschia sp. P. aminophilus P. gallaeciensis P. methylutens R. litoralis R. sphaeroides R. sphaeroides	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL	NC.015176 NC.004968 NC.005208 NC.005208 NC.007165 NC.011073 57 stability m NC.010905 NC.018022 NC.018022 NC.018022 NC.018022 NC.018022 NC.018022 NC.018023 NC.006362 NC.0103714 NC.004854 NC.011150 NC.011151 NC.007486 NC.012523 NC.013420 NC.013420 NC.013420 NC.013775 58 stability m NC.003078 NC.017323 NC.017323 NC.017323 NC.017326 NC.017323 NC.017326 NC.017323 NC.017326 NC.017323 NC.017326 NC.017323 NC.017326 NC.017326 NC.017323 NC.017326 NC.017323 NC.017326 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755 NC.01755	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi S. truber M. fulvus P. damselae measure : 0.420319 S. meliloti S. moliloti S. moliloti S. moliloti S. moliloti S. moliloti S. moliloti S. moliloti S. meliloti S. meliloti S. meliloti S. meliloti S. meliloti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria Aproteobacteria Chroobacteria Chroobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.56 0.58 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.18 nan 0.69 1.00 1.	PL PL PL PL PL PL CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009926 NC.009928 NC.009929 NC.009930 NC.009930 NC.009930 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012556 NC.009838 NC.01365 NC.01365 NC.01365 NC.01365 NC.00384 NC.00384 NC.00384 NC.009981 NC.016825 NC.019114 NC.004989 NC.005211 63 stability m NC.00384 NC.007801 NC.017304 64 stability m NC.012983 NC.018422 NC.018422 NC.018422 NC.018729 NC.007490 NC.007489 NC.007489 NC.007489 NC.007489 NC.007489 NC.007489 NC.007489	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. coli K. pneumoniae K. pneumoniae K. pneumoniae S. enterica S. marcescens Marcescens Marcescens Marcescens Marcescens Marcescens P. marcescens Marcescens P. aminophilus P. gallaeciensis P. methylutens R. sphaeroides R. sphaeroides R. denitrificans	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_01073 57 stability m NC_010905 NC_018580 NC_008703 NC_008703 NC_008703 NC_006362 NC_01314 NC_002576 NC_004854 NC_011150 NC_011150 NC_01151 NC_004854 NC_011523 NC_01459 NC_01459 NC_01372 NC_013775 58 stability m NC_009620 NC_017326 NC_017326 NC_017326 NC_017326 NC_017326 NC_019428 NC_010480 NC_010480 NC_010480 NC_012527	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense N. farcinica P. dioxanivorans R. equi R. equi S. melioti S. meliloti S. meliloti D. deserti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria aproteobacteria Aproteobacteria Chroobacteria Deinococci	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.56 0.38 0.64 0.64 0.64 1.00 0.82 nan 0.69 1.000 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012556 NC.008388 NC.013865 NC.008388 NC.013865 NC.009818 NC.009818 NC.0098114 NC.009912 NC.005211 63 stability m NC.009012 NC.017304 64 stability m NC.009956 NC.012983 NC.017304 64 stability m NC.009551 NC.013513 NC.018288 NC.018729 NC.00753 NC.00753 NC.007490 NC.007489 NC.008387 NC.008387 NC.008387 NC.008387	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica Measure : 0.919630 E. cloacae E. cloacae E. cloacae E. coli E. coli K. pneumoniae S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. enterica S. marcescens Marc	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γ proteobacteria γ proteobacteria α proteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL	NC.015176 NC.004968 NC.005208 NC.005208 NC.007165 NC.011073 57 stability m NC.010905 NC.018022 NC.018022 NC.018023 NC.006362 NC.004802 NC.018023 NC.006362 NC.018023 NC.006362 NC.018023 NC.004854 NC.01150 NC.01150 NC.014159 NC.014159 NC.010372 NC.01372 NC.01372 NC.013726 NC.017326 NC.017326 NC.014504 NC.012527 NC.012527 NC.012527 NC.012527 NC.012528	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense M. chubuense N. farcinica P. dioxanivorans R. equi R. equi S. meliolis S. meliloti S. meliloti D. deserti D. deserti D. deserti D. deserti	Bacilli Bacteria Bacilli Bacilli Bacilli Bacteria Bacilli Bacilli Bacilli Bacilli Bacteria Bacilli Bacilli Bacilli Bacteria Bach Bach Bach Bach Bach Bach Bach Bac	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.50 0.90 0.68 0.57 0.38 0.64 0.64 0.64 1.00 0.82 nan 0.82 nan 0.82 nan 0.69 1.00 1.0	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.009926 NC.009928 NC.009930 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.013365 NC.009838 NC.013365 NC.013365 NC.003384 NC.009980 NC.003384 NC.003848 NC.009981 NC.0099114 NC.009912 NC.017304 : 64 stability m NC.009912 NC.017304 : 64 stability m NC.009956 NC.013513 NC.018288 NC.018288 NC.018288 NC.018288 NC.018729 NC.007490 NC.007489 NC.007489 NC.007489 NC.008387 NC.008387 NC.008389 NC.009007	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica Cyanothece sp. S. baltica S. baltica S. baltica E. colo E. cloacae E. coli K. pneumoniae K. pneumoniae S. enterica S. marcescens marcescens Marcescens D. shibae H. baltica Jannaschia sp. P. adlaeciensis P. gallaeciensis P. gallaeciensis P. methylutens R. litoralis R. sphaeroides R. denitrificans R. sphaeroides R. denitrificans R. sphaeroides R. sphaeroides	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0
PL PL	NC_015176 NC_004968 NC_005208 NC_005208 NC_007165 NC_011073 57 stability m NC_010905 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_018022 NC_018023 NC_006362 NC_004854 NC_004854 NC_011150 NC_014159 NC_01320 NC_013420 NC_013420 NC_013420 NC_01375 58 stability m NC_009620 NC_003078 NC_01326 NC_01323 NC_017326 NC_012527 NC_012528 NC_012528 NC_012528	S. simulans S. thermophilus S. warneri S. warneri B. fragilis measure : 0.476968 A. benzoatilytica Gordonia sp. K. radiotolerans M. chubuense Mycobacterium sp. N. farcinica P. dioxanivorans R. equi R. equi S. meliolis S. meliloti S. meliloti S. meliloti S. meliloti Mathematica S. Synchococcus sp. D. deserti D. deserti D. deserti D. deserti D. deserti D. deserti	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacteroidia Actinobacteridae Actinobacteria aproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria Chroobacteria Chroobacteria Deinococci Deinococci	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.69 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.54 0.50 0.90 0.68 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.69 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.54 0.54 0.54 0.54 0.54 0.54 0.57 0.38 0.64 0.64 1.00 0.82 nan 0.69 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.009926 NC.009928 NC.009929 NC.009930 NC.009930 NC.009933 NC.011885 NC.009052 NC.017579 NC.017579 NC.012555 NC.012555 NC.012555 NC.012556 NC.009838 NC.013365 NC.013365 NC.013365 NC.013365 NC.013826 NC.003384 NC.003844 NC.003981 NC.016825 NC.01311 NC.004989 NC.004989 NC.004989 NC.004921 Stability m NC.009956 NC.012983 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018288 NC.018289 NC.007490 NC.007490 NC.007490 NC.007490 NC.008387 NC.008389 NC.009007 NC.009430	A. marina A. marina A. marina A. marina A. marina A. marina Cyanothece sp. S. baltica S. baltica S. baltica S. baltica E. cloacae E. cloacae E. cloacae E. cloi E. coli K. pneumoniae S. enterica S. marcescens E. coli M. pneumoniae S. enterica S. enterica S. enterica S. marcescens D. shibae H. baltica Jannaschia sp. P. aminophilus P. gallaeciensis P. gallaeciensis P. methylutens R. kitoralis R. sphaeroides R. denitrificans R. sphaeroides R. sphaeroides	Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria αproteobacteria	0.95 0.97 1.00 1.00 0.93 0.96 0.79 0.21 0.63 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0

267

Annexe D.Résultats du clustering de V^R par INFOMAP

PI.	NC 009040	B sphaeroides	oproteobacteria	1.00	PI.	NC 009128	Corunebacterium sp	Actinobacteridae	1 1 00
PL	NC_011960	R. sphaeroides	aproteobacteria	1.00	PL	NC_009129	Corvnebacterium sp.	Actinobacteridae	1.00
\mathbf{PL}	NZ_CM00116	3 R. sphaeroides	α proteobacteria	1.00	\mathbf{PL}	NC_005016	M. avium	Actinobacteridae	1.00
\mathbf{PL}	NC_004929	Ruegeria sp.	α proteobacteria	1.00	\mathbf{PL}	NC_009340	M. gilvum	Actinobacteridae	1.00
\mathbf{PL}	NC_008042	Ruegeria sp.	α proteobacteria	1.00	\mathbf{PL}	NC_014812	M. gilvum	Actinobacteridae	1.00
\mathbf{PL}	NC_016000	S. chungbukense	α proteobacteria	0.73	\mathbf{PL}	NC_{005206}	T. pyogenes	Actinobacteridae	1.00
\mathbf{PL}	NC_017965	A. kashmirensis	β proteobacteria	1.00	\mathbf{PL}	NC_017792	D. gobiensis	Deinococci	0.88
PL	NC_006628	L. hongkongensis	β proteobacteria	0.50					
PL	NC_007616	N. multiformis	β proteobacteria	1.00	cluster	: 70 stability	measure : 0.724444		
PL	NC_008762	P. naphthalenivorans	β proteobacteria	0.88	CHR	NC_015185	D. thermolithotrophum	Aquificae	0.77
PL	NC_013962	Cand. Riesia	γ proteobacteria	1.00	CHR	NC_012440	P. marina	Aquificae	0.88
.1	05	1 000000			CHR	NC_012438	S. azorense	Aquificae	0.70
DI	NC 011017	Eli		1.00	CHR	NC_010730	Sulfurihydrogenibium sp.	Aquificae	0.88
PL DI	NC 018651	E. coli	~proteobacteria	1.00	Спк	NC_014920	1. ammonificans	Aquincae	0.57
PL	NC 010152	K preumoniae	oproteobacteria	1.00	clustor	• 71 stability	$m_{02} = 0.061905$		
PL	NC 016838	K pneumoniae	vproteobacteria	1.00	CHB	NC 009497	M analactine	Mollicutes	1.00
PL	NC_003385	S. enterica	γ proteobacteria	1.00	CHR	NC 013948	M. agalactiae	Mollicutes	1.00
PL	NC_003134	Y. pestis	γ proteobacteria	1.00	CHR	NC_014921	M. fermentans	Mollicutes	1.00
\mathbf{PL}	NC_004835	Y. pestis	γ proteobacteria	1.00	CHR	NC_014552	M. fermentans	Mollicutes	1.00
\mathbf{PL}	NC_004838	Y. pestis	γ proteobacteria	1.00	CHR	NC_014760	M. bovis	Mollicutes	1.00
\mathbf{PL}	NC_005815	Y. pestis	γ proteobacteria	1.00	CHR	NC_015725	M. bovis	Mollicutes	1.00
\mathbf{PL}	NC_006323	Y. pestis	γ proteobacteria	1.00	CHR	NC_018077	M. bovis	Mollicutes	1.00
PL	NC_009596	Y. pestis	γ proteobacteria	1.00	\mathbf{PL}	NC_007386	S. citri	Mollicutes	1.00
\mathbf{PL}	NC_009378	Y. pestis	γ proteobacteria	1.00					
PL	NC_008118	Y. pestis	γ proteobacteria	1.00	cluster	: 72 stability	measure : 0.470243		
PL	NC_008120	Y. pestis	γ proteobacteria	1.00	PL	NC_014917	T. ammonificans	Aquificae	0.43
PL	NC_010158	Y. pestis	γproteobacteria	1.00	CHR	NC_013943	D. acetiphilus	Deferribacteres	0.62
PL DI	NC 017266	I. pestis	~proteobacteria	1.00	CHR	NC_014216	D. alkaliphilus	oproteobacteria	0.73
PL	NC 017155	1. pestis V pestis	oproteobacteria	1.00	CHR	NC_014972	D. propionicus	oproteobacteria	0.80
PL	NC 017158	V nestis	oproteobacteria	1.00	Спк	NC_000138	D. psychrophila	oproteobacteria	0.80
PL	NC_017169	Y. pestis	vproteobacteria	1.00	cluster	• 73 stability	measure : 0.746389		
	110-011100	11 peoble	/processaeceria	1100	CHR	NC 015174	P brasiliansis	Planctomycotacia	1.00
cluster	: 66 stability	measure : 0.989474			CHB	NC 014148	P limnonhilus	Planctomycetacia	1.00
PL	NC_008568	B. afzelii	Spirochaetes	1.00	CHR	NC_013720	P. stalevi	Planctomycetacia	0.77
\mathbf{PL}	NC_011789	B. afzelii	Spirochaetes	1.00	CHR	NC_005027	R. baltica	Planctomycetacia	0.73
\mathbf{PL}	NC_017232	B. afzelii	Spirochaetes	1.00				v	1
\mathbf{PL}	NC_015915	B. bissettii	Spirochaetes	1.00	cluster	: 74 stability	measure : 0.651140		
PL	NC_004988	B. burgdorferi	Spirochaetes	1.00	\mathbf{PL}	NC_013546	A. pleuropneumoniae	γ proteobacteria	1.00
PL	NC_001849	B. burgdorferi	Spirochaetes	1.00	\mathbf{PL}	NC_013718	$C. \ rodentium$	γ proteobacteria	0.07
PL	NC_011865	B. burgdorferi	Spirochaetes	1.00	\mathbf{PL}	NC_005706	E. amylovora	γ proteobacteria	1.00
PL	NC_012188	B. burgdorferi	Spirochaetes	1.00	PL	NC_013957	E. amylovora	γ proteobacteria	1.00
PL	NC_012227	B. burgaorferi	Spirochaetes	1.00	PL	NC_013972	E. amylovora	γ proteobacteria	1.00
PL	NC 012237	B. burgaorferi B. burgaorferi	Spirochaetes	1.00	PL	NC_018978	E. amylovora	γ proteobacteria	1.00
	NC 012170	B. burgdonfeni	Spirochaetes	1.00	PL	NC_011351	E. coli	γ proteobacteria	0.30
PL	NC 012179	B. burgdorferi	Spirochaetes	1.00	PL	NC_017002	E. coli Eli	γproteobacteria	0.41
PL	NC 017410	B. burgdorferi	Spirochaetes	1.00		NC 004445	E. coli E. partifolian	γproteobacteria	1.00
PL	NC_017417	B. burgdorferi	Spirochaetes	1.00	PL	NC 013263	E. pyrijonae E. pyrijonae	oproteobacteria	1.00
\mathbf{PL}	NC_018992	B. burgdorferi	Spirochaetes	1.00	PL	NC 017392	E nurifoliae	vproteobacteria	1.00
\mathbf{PL}	NC_011861	B. garinii	Spirochaetes	1.00	PL	NC_010696	E. tasmaniensis	γproteobacteria	0.67
\mathbf{PL}	NC_011871	B. garinii	Spirochaetes	1.00	\mathbf{PL}	NC_004834	Erwinia sp.	γ proteobacteria	1.00
\mathbf{PL}	NC_012226	B. hyodysenteriae	Spirochaetes	1.00	\mathbf{PL}	NC_017442	Erwinia sp.	γ proteobacteria	1.00
PL	NC_012200	B. spielmanii	Spirochaetes	1.00	\mathbf{PL}	NC_004771	P. multocida	γ proteobacteria	1.00
\mathbf{PL}	NC_012180	B. valaisiana	Spirochaetes	1.00	\mathbf{PL}	NC_016835	R. aquatilis	γ proteobacteria	0.42
PL	NC_012242	Borrelia sp.	Spirochaetes	1.00	\mathbf{PL}	NC_017060	R. aquatilis	γ proteobacteria	0.36
PL	NC_015386	T. succinifaciens	Spirochaetes	1.00	\mathbf{PL}	NC_015062	Rahnella sp.	γ proteobacteria	0.36
.1	07	0.200100			PL	NC_010657	S. boydii	γ proteobacteria	0.05
DI	NC 016007	C = abvises = = = = = = = = = = = = = = = = = =	A stimely stanides	0.47	PL	NC_011148	S. enterica	γ proteobacteria	1.00
CUD	NC 014255	G. polyisoprenivorans	Nitroaping	0.47	PL	NC_019128	S. enterica	γ proteobacteria	0.30
CHR	NC 018649	L. ferrinhilum	Nitrospira	0.03	PL	NC_019130	S. enterica	γ proteobacteria	0.12
CHR	NC_017094	L. ferrooxidans	Nitrospira	0.36	cluster	• 75 stability	measure : $1,000000$		
CHR	NC_015318	H. maritima	δ proteobacteria	1.00	CHB	NC 012225	B hyodusenteriae	Spirochaetes	1.00
	•		-		CHR	NC_017243	B. intermedia	Spirochaetes	1.00
cluster	: 68 stability	measure : 1.000000			CHR	NC_014150	B. murdochii	Spirochaetes	1.00
CHR	NC_000918	A. aeolicus	Aquificae	1.00	\mathbf{CHR}	NC_014330	B. pilosicoli	Spirochaetes	1.00
CHR	NC_013799	H. thermophilus	Aquificae	1.00	CHR	NC_018607	B. pilosicoli	Spirochaetes	1.00
CHR	NC_017161	H. thermophilus	Aquificae	1.00					
CHR	NC_011126	Hydrogenobaculum sp.	Aquificae	1.00	cluster	: 76 stability	measure : 0.650942		
CHR	NC_013894	T. albus	Aquificae	1.00	PL	NC_011561	Cand. Azobacteroides	Bacteroidia	0.62
aluston	. 60 stability	monguro (0.020500			PL	NC_011562	Cand. Azobacteroides	Bacteroidia	0.94
DI	NC 010813	C aurimucosum	Actinobactoridae	1.00	PL	NC_018742	E. oligotrophica	Cytophagia	1.00
PL	NC 003239	C. diphtheriae	Actinobacteridae	1.00	PL PL	NC 012705	D magneticus	Aproteobacteria	0.70
PL	NC_005001	C. diphtheriae	Actinobacteridae	1.00	PL	NC 006134	C coli	eproteobacteria	1 00
PL	NC_004319	C. efficiens	Actinobacteridae	1.00	PL	NC_006135	C. jejuni	eproteobacteria	1.00
\mathbf{PL}	NC_004320	C. efficiens	Actinobacteridae	1.00	PL	NC_007141	C. jejuni	εproteobacteria	1.00
\mathbf{PL}	NC_001456	C. glutamicum	Actinobacteridae	1.00	\mathbf{PL}	NC_008790	C. jejuni	ϵ proteobacteria	1.00
\mathbf{PL}	NC_003227	C. glutamicum	Actinobacteridae	1.00	\mathbf{PL}	NC_014801	C. jejuni	ϵ proteobacteria	1.00
\mathbf{PL}	NC_004945	C.~glutamicum	Actinobacteridae	1.00	\mathbf{PL}	NC_{017282}	C. jejuni	ϵ proteobacteria	1.00
\mathbf{PL}	NC_001415	$C. \ glutamicum$	Actinobacteridae	1.00	\mathbf{PL}	$NC_{-}012040$	C. lari	ϵ proteobacteria	1.00
PL	NC_013726	$C. \ glutamicum$	Actinobacteridae	1.00	\mathbf{PL}	NC_{014754}	S. kujiense	ϵ proteobacteria	0.00
PL	NC_002775	C. jeikeium	Actinobacteridae	1.00	\mathbf{PL}	NC_{009777}	V. harveyi	γ proteobacteria	0.70
PL	NC_003080	C. jeikeium	Actinobacteridae	1.00	-				
PL DT	NC_003442	C. jeikeium	Actinobacteridae	1.00	cluster	: 77 stability	measure : 0.945657	<u> </u>	1
PL DI	NC 004772	C. jeikeium	Actinobacteridae	1.00	PL	NC_010543	Cyanothece sp.	Chroobacteria	1.00
г L рт	NC 004774	C. jeikeium	Actinobacteridae	1.00	PL	NC_002145	E. coli	γ proteobacteria	1.00
PL	NC.004953	C. jeikeium	Actinobacteridae	1.00	PL DI	NC_017620	E. coli	γproteopacteria	1.00
PL	NC_014167	C. resistens	Actinobacteridae	0.00	PL PL	NC 017650	E. coli	vproteobacteria	1.00
PL	NC_004939	C. striatum	Actinobacteridae	0.33	PL	NC 018652	E. coli	vproteobacteria	1 00
				1				/	1 1.00

БТ	L NC 019660	E li		1 1 00
PL	NC_018660	E. coli	γ proteobacteria	1.00
PL	NC_018663	E. coli	γproteobacteria	1.00
PL	NC_019059	E. coli	γ proteobacteria	1.00
PL	NC_019075	E. coli	γ proteobacteria	1.00
\mathbf{PL}	NC_004345	H. somni	γ proteobacteria	1.00
PL.	NC 016841	K nneumoniae	vproteobacteria	1.00
	1	<i>F</i>	/	
clustor	• 78 etability m	0.25000 ± 0.728828		
Cluster	NG 012005	A 1 11 1	D	1.00
CHR	NC_013205	A. aciaocalaarius	Bacilli	1.00
PL	NC_013206	A. acidocaldarius	Bacilli	0.93
PL	NC_013207	A. acidocaldarius	Bacilli	1.00
CHR	NC_017167	A. acidocaldarius	Bacilli	1.00
cluster	: 79 stability m	neasure : 0.812996		
PL	NC 007682	E coli	vproteobacteria	1.00
PL	NC 009132	E. coli	oproteobacteria	1.00
FL	NC_009132	E. con	proteobacteria	1.00
PL	NC_014231	E. coli	γproteobacteria	1.00
PL	NC_015599	E. coli	γ proteobacteria	1.00
\mathbf{PL}	NC_019033	E. coli	γ proteobacteria	1.00
\mathbf{PL}	NC_019082	E. coli	γ proteobacteria	1.00
\mathbf{PL}	NC_019087	E. coli	γ proteobacteria	1.00
PL	NC_019098	E. coli	γ proteobacteria	1.00
PL.	NC 018999	E amulovora	vproteobacteria	1.00
DI	NC 011282	K moumoniae	aprotosbacteria	1.00
PL	NC 011365	K. pneumoniae	γproteobacteria	1.00
PL	NC_011385	K. pneumoniae	γ proteobacteria	1.00
PL	NC_011617	K. pneumoniae	γ proteobacteria	1.00
PL	NC_014368	K. pneumoniae	γ proteobacteria	1.00
PL	NC_019166	K. pneumoniae	γ proteobacteria	1.00
\mathbf{PL}	NC_014208	K. oxytoca	γ proteobacteria	1.00
\mathbf{PL}	NC_009980	S. enterica	γ proteobacteria	1.00
\mathbf{PI}	NC_019124	S. enterica	vproteobacteria	1.00
	1		/ P	1
clustor	• 80 etability m	0.035000 ± 0.003056		
DI	NG 000050	D_1:1-		1.00
PL	NC_009959	D. shibae	α proteobacteria	1.00
PL	NC_014621	K. vulgare	α proteobacteria	1.00
\mathbf{PL}	NC_017386	K. vulgare	α proteobacteria	1.00
\mathbf{PL}	NC_018423	P. gallaeciensis	α proteobacteria	1.00
\mathbf{PL}	NC_018287	P. gallaeciensis	α proteobacteria	1.00
RECE	NC_007494	R. sphaeroides	α proteobacteria	1.00
RECE	NC_009050	R. sphaeroides	α proteobacteria	1.00
RECE	NC_009429	R. sphaeroides	α proteobacteria	1.00
RECE	NC 011958	R. sphaeroides	αproteobacteria	1.00
PL.	NC 008043	Buegeria SD	aproteobacteria	0.54
	1 110 20000 10	Huogon tu opi	aprotobactoria	0.01
cluster	• 81 stability m	easure : 0.842979		
DI	i oʻr stasintij n	iododi o i olo izolo		
	NC 005328	B methanolicus	Bacilli	1 1 00
PL	NC_005328	B. methanolicus	Bacilli	1.00
PL	NC_005328 NC_006509	B. methanolicus G. kaustophilus	Bacilli Bacilli	1.00
PL PL PL	NC_005328 NC_006509 NC_010420	B. methanolicus G. kaustophilus G. stearothermophilus	Bacilli Bacilli Bacilli	1.00 1.00 1.00
PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius	Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00
PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp.	Bacilli Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_013412	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00 \end{array} $
PL PL PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_013412 NC_014651	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array} $
PL PL PL PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_013412 NC_014651 NC_014916	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$ \begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array} $
PL PL PL PL PL PL PL PL	NC.005328 NC.006509 NC.010420 NC.015665 NC.012794 NC.013412 NC.014651 NC.014916 NC.009700	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia	$ \begin{array}{c} 1.00\\ 1.00$
PL PL PL PL PL PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_013412 NC_014651 NC_014916 NC_009700 NC_017298	B. methanolicus G. kaustophilus G. stearothermophilus Geobacillus sp. Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_013412 NC_014916 NC_009700 NC_017298	B. methanolicus G. kaustophilus G. stearothermophilus Geobacillus sp. Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. olicatraphica	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cutanbaria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL	NC.005328 NC.006509 NC.010420 NC.015665 NC.012794 NC.013412 NC.014651 NC.014916 NC.009700 NC.017298 NC.018743	B. methanolicus G. kaustophilus G. stearothermophilus Geobacillus sp. Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_013651 NC_014651 NC_014651 NC_0146916 NC_009700 NC_018749 NC_018749 NC_018749	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 nan 1.00 0.43 0.02
PL PL PL PL PL PL PL PL PL PL PL	NC.005328 NC.006509 NC.010420 NC.012794 NC.013412 NC.014916 NC.004916 NC.014916 NC.014916 NC.017298 NC.017298 NC.018749 NC.002806	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 nan 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL	NC.005328 NC.006509 NC.010420 NC.015665 NC.012794 NC.014651 NC.014916 NC.009700 NC.017298 NC.018743 NC.018743 NC.018749 NC.002806 NC.015705	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_014511 NC_014651 NC_014651 NC_014916 NC_009700 NC_017298 NC_018749 NC_018749 NC_018749 NC_015705 NC_013737	B. methanolicus G. kaustophilus G. stearothermophilus Geobacillus sp. Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00
FL PL PL PL PL PL PL PL PL PL PL PL PL	NC.005328 NC.006509 NC.010420 NC.012794 NC.013412 NC.014651 NC.014916 NC.009700 NC.017298 NC.018749 NC.018749 NC.008749 NC.018705 NC.013738	B. methanolicus G. kaustophilus G. stearothermophilus Geobacillus sp. Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.015665 NC.012794 NC.014651 NC.014916 NC.009700 NC.017298 NC.018749 NC.002806 NC.018749 NC.002806 NC.013737 NC.013737 NC.013738 NC.007678	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_014651 NC_014651 NC_014651 NC_014674 NC_014743 NC_018743 NC_018743 NC_018749 NC_018745 NC_015705 NC_013737 NC_013738 NC_016936	B. methanolicus G. kaustophilus G. stearothermophilus Geobacillus sp. Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_005328 NC_006509 NC_010420 NC_010420 NC_012794 NC_014651 NC_014651 NC_014651 NC_014651 NC_014651 NC_014651 NC_014739 NC_018749 NC_008740 NC_018749 NC_002806 NC_013737 NC_013737 NC_013738 NC_007678 NC_006140	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.95
FL PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.005328 NC.006509 NC.010420 NC.012794 NC.013412 NC.014916 NC.014916 NC.014916 NC.014916 NC.017298 NC.018749 NC.018749 NC.018749 NC.018749 NC.018738 NC.013738 NC.013738 NC.016936 NC.016936 NC.012733	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria &proteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_005328 NC_006509 NC_010420 NC_015665 NC_012794 NC_014651 NC_014651 NC_014651 NC_014674 NC_018743 NC_018743 NC_018743 NC_018749 NC_018749 NC_015705 NC_013737 NC_013737 NC_013738 NC_016936 NC_006140 NC_012733	 B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum C. botulinum E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri 	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.95\\ 1.00\\ \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.015665 NC.012794 NC.01451 NC.014651 NC.014651 NC.014651 NC.014798 NC.018743 NC.018743 NC.018749 NC.0018749 NC.018749 NC.018775 NC.013737 NC.013738 NC.013737 NC.013738 NC.007678 NC.006140 NC.012733 : 82 stability m	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria	$\begin{array}{c} 1.00\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_005328 NC_006509 NC_010420 NC_010420 NC_012794 NC_013412 NC_014651 NC_009700 NC_017298 NC_018743 NC_018743 NC_018749 NC_002806 NC_018705 NC_013737 NC_013738 NC_016936 NC_007678 NC_016936 NC_006140 NC_002733	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri easure : 0.418822	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria &proteobacteria #proteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.015665 NC.013412 NC.014916 NC.014916 NC.017298 NC.018749 NC.018749 NC.018749 NC.018738 NC.013738 NC.0107678 NC.012733 * 82 stability m NC.009718 NC.009718	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum E. ocosum	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria cproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.012794 NC.014651 NC.014916 NC.001708 NC.018749 NC.018749 NC.018737 NC.013737 NC.013738 NC.007678 NC.006140 NC.012733 s2 stability m NC.009718 NC.014529	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria cproteobacteria cproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013412 NC.014651 NC.01916 NC.009700 NC.017298 NC.018743 NC.018743 NC.018743 NC.018749 NC.002806 NC.018737 NC.013737 NC.013737 NC.013738 NC.016936 NC.006140 NC.012733 : 82 stability m NC.009718 NC.017095 NC.011653	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria εproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.012794 NC.013412 NC.014916 NC.014916 NC.017298 NC.018749 NC.018749 NC.013738 NC.013738 NC.016936 NC.012733 82 stability m NC.00705 NC.01705 NC.01653 NC.01653	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria froteobacteria Thermotogae Thermotogae Thermotogae	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_005328 NC_005328 NC_010420 NC_010420 NC_010420 NC_010420 NC_010420 NC_015665 NC_014916 NC_014916 NC_009700 NC_017298 NC_018743 NC_018743 NC_018705 NC_013737 NC_013738 NC_006160 NC_016936 NC_0101703 82 stability m NC_00718 NC_017095 NC_01653 NC_009616	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria ϵproteobacteria thermotogae Thermotogae Thermotogae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 0.43 1.00 0.85 1.00 0.88 0.88
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.012794 NC.014651 NC.014916 NC.001709 NC.018743 NC.018749 NC.002806 NC.013737 NC.013738 NC.001636 NC.0012733 82 stability m NC.009718 NC.009616 83 stability m	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria &proteobacteria terratoria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_005328 NC_006509 NC_010420 NC_012794 NC_013412 NC_014916 NC_017298 NC_018749 NC_018749 NC_013737 NC_016936 NC_016936 NC_017338 NC_007678 NC_001705 NC_012733 * 82 stability n NC_01705 NC_01705 NC_01705 NC_01653 NC_009616 * 83 stability n	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria εproteobacteria thermotogae Thermotogae Thermotogae	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013412 NC.014651 NC.014916 NC.009700 NC.017298 NC.018749 NC.008749 NC.018749 NC.002806 NC.015705 NC.013738 NC.016336 NC.016376 NC.012733 : 82 stability m NC.017095 NC.011653 NC.017095 NC.011653 NC.009616 : 83 stability m NC.010541 NC.010547	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pensivorans T. africanus T. melanesiensis	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria δproteobacteria Thermotogae Thermotogae Thermotogae	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.01420 NC.013731 NC.018743 NC.018743 NC.018743 NC.018743 NC.018775 NC.013737 NC.016936 NC.016936 NC.016936 NC.016936 NC.017095 NC.016936 NC.017095 NC.01653 NC.01653 NC.009616 : 83 stability n NC.010541 NC.010541	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria eproteobacteria Thermotogae Thermotogae Thermotogae Cheroobacteria Chroobacteria	1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 0.73 0.73 0.88 0.73
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.012794 NC.014651 NC.014651 NC.014651 NC.014651 NC.01798 NC.018749 NC.018749 NC.018749 NC.018737 NC.013737 NC.013738 NC.006140 NC.012733 82 stability m NC.009718 NC.009718 NC.009616 : 83 stability m NC.010541 NC.010547 NC.011721 NC.011723	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria \$proteobacteria eproteobacteria Thermotogae Thermotogae Thermotogae Chroobacteria Chroobacteria Chroobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 1.00 0.43 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.014651 NC.014916 NC.009700 NC.017298 NC.018749 NC.008749 NC.018749 NC.0018749 NC.018749 NC.018749 NC.018749 NC.018743 NC.013738 NC.013738 NC.013738 NC.013738 NC.016936 NC.010768 NC.016936 NC.010705 NC.011653 NC.017095 NC.011653 NC.009616 : 83 stability m NC.010541 NC.010547 NC.011723 NC.011723	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri Heasure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis Heasure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Choobacteria Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria Chroobacteria Chroobacteria	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.01420 NC.01376 NC.018743 NC.018743 NC.018743 NC.018743 NC.0187743 NC.013737 NC.013738 NC.0016705 NC.016936 NC.0016936 NC.016936 NC.016936 NC.016936 NC.016936 NC.016936 NC.01653 NC.01653 NC.01653 NC.016541 NC.010541 NC.011721 NC.011723 NC.011738	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp. Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria cproteobacteria Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria Chroobacteria Chroobacteria Chroobacteria	1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 1.00 1.00 1.00 0.43 1.00 1.05 1.55 1
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.012794 NC.014651 NC.014651 NC.014651 NC.014916 NC.001709 NC.018743 NC.018749 NC.002806 NC.013737 NC.013738 NC.0013738 NC.006140 NC.012733 82 stability m NC.009718 NC.009718 NC.010705 NC.010531 NC.009616 : 83 stability m NC.010541 NC.010541 NC.011721 NC.011737 NC.011738 NC.011737 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Chroebacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	1.00 1.05 1.00 1.05 1
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.014916 NC.014916 NC.009700 NC.017298 NC.018749 NC.018749 NC.008749 NC.018749 NC.018749 NC.018749 NC.018743 NC.018743 NC.013738 NC.013738 NC.013738 NC.016936 NC.016936 NC.016936 NC.016936 NC.016936 NC.01705 NC.011653 NC.017095 NC.011653 NC.009616 : 83 stability m NC.010541 NC.010541 NC.011721 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri Measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Choobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013665 NC.012794 NC.014651 NC.014651 NC.014651 NC.014651 NC.018749 NC.018749 NC.018749 NC.018749 NC.018749 NC.013737 NC.013738 NC.013738 NC.016936 NC.016936 NC.016936 NC.016936 NC.016936 NC.016936 NC.01653 NC.01653 NC.016541 NC.010541 NC.01721 NC.011721 NC.011723 NC.011738 NC.011738 NC.011738 NC.011882 NC.011882 NC.011860	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria &proteobacteria Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013665 NC.012794 NC.014651 NC.014651 NC.014651 NC.014651 NC.014651 NC.018749 NC.008749 NC.018749 NC.002806 NC.018749 NC.013737 NC.013737 NC.013737 NC.013738 NC.007678 NC.006140 NC.0012733 : 82 stability m NC.009718 NC.0106541 NC.0107541 NC.011721 NC.011721 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011882 NC.011882 NC.01360 NC.014502	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum C. botulinum E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothec	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria oproteobacteria Thermotogae Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria	1.00 0.73 0.73 0.88 0.77 0.73 0.73 0.77 0.77 0.53 0.82 1.00 0.55 1.00 0.77 0.53 0.82 1.00 0.55 0.00 0.55 0.77 0.57 0.55 0.82 1.00 0.55 0.00 0.55 0
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.014651 NC.012794 NC.014651 NC.017298 NC.018749 NC.008740 NC.018749 NC.008743 NC.018749 NC.018749 NC.013737 NC.013737 NC.013737 NC.013738 NC.007678 NC.016936 NC.006936 NC.006936 NC.0016036 NC.017095 NC.011653 NC.017095 NC.011653 NC.009616 : 83 stability m NC.010541 NC.010723 NC.011737 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis easure : 0.635224 Cyanothece sp. Cyanothece	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Choobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	1.00 0.73 0.88 0.73 0.73 0.53 0.77 0.53 0.50 1.00 1.00 0.50 0.77 0.55
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013665 NC.012794 NC.014651 NC.014651 NC.014651 NC.014703 NC.018743 NC.018743 NC.018743 NC.018743 NC.018743 NC.018743 NC.013737 NC.013738 NC.013737 NC.013738 NC.016936 NC.016936 NC.016936 NC.016936 NC.01653 NC.01721 NC.01721 NC.01721 NC.01721 NC.01723 NC.011738 NC.011738 NC.011738 NC.011882 NC.014502 NC.014533 NC.004073 NC.014533 NC.004073	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. C	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria δproteobacteria chroteobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	1.00 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.55 0.55 0.55 0.50 1.00 1.00 0.75 0.55
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.005328 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.012794 NC.014651 NC.01798 NC.018743 NC.018743 NC.018749 NC.002806 NC.013737 NC.013738 NC.0013738 NC.006140 NC.012733 S2 stability m NC.009718 NC.009718 NC.009616 : 83 stability m NC.010541 NC.010541 NC.011721 NC.011738 NC.011737 NC.011738 NC.014503 NC.014503 NC.004073 NC.0047595 <th>B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri Measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothe</th> <th>Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Choobacteria Thermotogae Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria</th> <th>1.00 0.73 0.73 0.88 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.55 0.73 0.55</th>	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri Measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothe	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Choobacteria Thermotogae Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria Chroobacteria	1.00 0.73 0.73 0.88 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.55 0.73 0.55
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013412 NC.014651 NC.012798 NC.01798 NC.01798 NC.018743 NC.018743 NC.018749 NC.002806 NC.018743 NC.018749 NC.018743 NC.013737 NC.013737 NC.013737 NC.013738 NC.007678 NC.0107678 NC.011733 NC.017095 NC.011653 NC.009616 : 83 stability m NC.010541 NC.010541 NC.0107541 NC.011723 NC.011737 NC.011738 NC.011737 NC.011738 NC.011738 NC.011738 NC.011738 NC.011737 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011735 NC.014502 NC.014502 NC.019440	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri T. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. S. elongatus S. elongatus Anabaena sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Cytophagia Choobacteria Chroobacteria	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013665 NC.012794 NC.01451 NC.014651 NC.014651 NC.014651 NC.018743 NC.018743 NC.018743 NC.018743 NC.018743 NC.013737 NC.013738 NC.013737 NC.013738 NC.016936 NC.016936 NC.016936 NC.016936 NC.01653 NC.001653 NC.011721 NC.011721 NC.011721 NC.011723 NC.011738 NC.011738 NC.011738 NC.011738 NC.011882 NC.0114502 NC.014503 NC.014503 NC.004630	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. S. elongatus Anabaena sp. N. punctiforme	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Sphingobacteria cproteobacteria proteobacteria Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria	1.00 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.55 0.55 0.50 0.00 0.00 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.50 0.00 0.00 0.00 0.50 0.50 0.50 0.00 0.00 0.00 0.00 0.00 0.00 0.50 0.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.01365 NC.014916 NC.009700 NC.01798 NC.018743 NC.018743 NC.018749 NC.002806 NC.013737 NC.013738 NC.0013738 NC.006140 NC.012733 Stability m NC.009718 NC.009718 NC.009616 Stastability m NC.010541 NC.010721 NC.011721 NC.011738 NC.011738 NC.011731 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.014533 NC.014533 NC.014533 NC.019440	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. Cyanothe	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Choebacteria Thermotogae Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria	1.00 0.73 0.73 0.73 0.73 0.73 0.73 0.73 0.77 0.550 0.00 0.90 0.50 0.50 0.50 0.00 0.00 0.50 0.50 0.00 0.00 0.00 0.50 0.50 0.00 0.00 0.00 0.50 0.50 0.00 0.00 0.00 0.00 0.50 0.50 0.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.014651 NC.012794 NC.014916 NC.009700 NC.017298 NC.018743 NC.018749 NC.002806 NC.018743 NC.018743 NC.018749 NC.002806 NC.013737 NC.013737 NC.013738 NC.013737 NC.013738 NC.0013738 NC.0013738 NC.001633 NC.0017095 NC.011723 NC.011723 NC.011723 NC.011723 NC.011723 NC.011737 NC.011738 NC.011738 NC.011738 NC.011738 NC.011737 NC.011738 NC.011738 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011735 NC.011735 NC.014503 NC.004073 NC.007595 NC.019440 NC.010633 NC.003270	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. ruber S. grandis D. psychrophila A. butzleri Measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Choobacteria Chroobacteria	$\begin{array}{c} 1.00\\ 0.95\\ 1.00\\ 1.00\\ 0.95\\ 1.00\\ 1.00\\ 0.50\\ 0.77\\ 0.77\\ 0.53\\ 0.88\\ 0.88\\ 0.88\\ 0.77\\ 0.77\\ 0.53\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.012794 NC.013665 NC.012794 NC.014916 NC.004651 NC.014916 NC.00700 NC.017298 NC.018743 NC.018743 NC.018743 NC.018743 NC.018743 NC.013737 NC.013737 NC.013738 NC.016936 NC.016936 NC.016936 NC.016936 NC.017095 NC.01721 NC.01721 NC.01721 NC.011721 NC.011723 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011733 NC.011738 NC.011733 NC.011738 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.011733 NC.014502 NC.014503 NC.004073 NC.007595 NC.019440 NC.010630 NC.003273	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. S. elongatus S. elongatus S. elongatus Anabaena sp. N. punctiforme N. punctiforme Nostoc sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Choebacteria Thermotogae Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria	$\begin{array}{c} 1.00\\ 0.95\\ 1.00\\ 0.95\\ 1.00\\ 0.73\\ 0.77\\ 0.73\\ 0.88\\ 0.88\\ 0.88\\ 0.77\\ 0.73\\ 0.73\\ 0.73\\ 0.73\\ 0.73\\ 0.73\\ 0.73\\ 0.73\\ 0.55\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.63\\ 0.63\\ 0.63\\ 0.00\\ 0.63\\ 0.00\\ 0.50\\ 0.63\\ 0.00\\ 0.50\\ 0.63\\ 0.00\\ 0.50\\ 0.63\\ 0.00\\ 0.50\\ 0.63\\ 0.00\\ 0.50\\ 0.50\\ 0.63\\ 0.00\\ 0.50\\ 0.50\\ 0.63\\ 0.50\\ 0.50\\ 0.50\\ 0.63\\ 0.50\\ 0.50\\ 0.63\\ 0.50\\ 0.50\\ 0.63\\ 0.50\\$
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC.005328 NC.006509 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.010420 NC.01420 NC.01420 NC.013731 NC.018743 NC.018743 NC.018743 NC.018743 NC.018749 NC.002806 NC.013737 NC.013738 NC.0013738 NC.006140 NC.012733 NC.006140 NC.010718 NC.009718 NC.009718 NC.009616 S3 stability m NC.011721 NC.011721 NC.011721 NC.011738 NC.011738 NC.011731 NC.011733 NC.011733 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738 NC.011738	B. methanolicus G. kaustophilus G. stearothermophilus G. thermoglucosidasius Geobacillus sp. Geobacillus sp. Geobacillus sp. C. botulinum E. oligotrophica E. oligotrophica Microscilla sp. R. slithyformis S. linguale S. linguale S. linguale S. linguale S. linguale S. ruber S. grandis D. psychrophila A. butzleri measure : 0.418822 F. nodosum F. pennivorans T. africanus T. melanesiensis measure : 0.635224 Cyanothece sp. Cyanothece sp. S. elongatus S. elongatus Anabaena sp. N. punctiforme N. punctiforme N. punctiforme Nostoc sp. Nostoc sp.	Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Cytophagia Choobacteria Thermotogae Thermotogae Thermotogae Thermotogae Thermotogae Thermotogae Chroobacteria	$\begin{array}{c} 1.00\\ 0.95\\ 1.00\\ 0.95\\ 1.00\\ 0.77\\ 0.73\\ 0.73\\ 0.73\\ 0.73\\ 0.73\\ 0.77\\ 0.53\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.88\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.50\\ 0.63\\ 0.68\\ 0.68\\ 0.68\\ 0.88\\ 0.68\\$

cluster : 84 stability measure : 0.976190

DI	NC 016022	Cli		1.00
	NC 004724	G. Lynnus	aproteobacteria	1.00
PL	NC_004734	A. calaus	γproteobacteria	1.00
PL	NC_015852	A. caldus	γ proteobacteria	1.00
\mathbf{PL}	NC_011207	A. hydrophila	γ proteobacteria	1.00
\mathbf{PL}	NC_005312	A. pleuropneumoniae	γ proteobacteria	1.00
\mathbf{PL}	NC_003123	A. salmonicida	γ proteobacteria	1.00
\mathbf{PL}	NC_003124	A. salmonicida	γ proteobacteria	1.00
\mathbf{PL}	NC_002636	D. nodosus	γ proteobacteria	1.00
PL.	NC 012006	E cloacae	oproteobacteria	1.00
PI	NC 017097	E cloacae	oproteobacteria	1.00
	NC 014256	E. cioacae	/proteobacteria	1.00
PL	NC_014556	E. coli	γproteobacteria	1.00
PL	NC_019091	E. coli	γproteobacteria	0.86
PL	NC_006994	P. multocida	γ proteobacteria	1.00
\mathbf{PL}	NC_011378	P. multocida	γ proteobacteria	1.00
\mathbf{PL}	NC_008764	P. naphthalenivorans	β proteobacteria	nan
\mathbf{PL}	NC_016859	S. enterica	γ proteobacteria	1.00
\mathbf{PL}	NC_016862	S. enterica	γ proteobacteria	1.00
\mathbf{PL}	NC 017719	S. enterica	vproteobacteria	1.00
DI	NC 012104	S. ontonica	protosbastoria	1.00
	NC 002524	D. enterica	Francestal	1.00
	NC_002524	uncultured	Environmental	1.00
PL	NC_004975	uncultured	Environmental	1.00
• .				
cluster	: 85 stability	measure : 1.000000		
CHR	NC_007633	M. capricolum	Mollicutes	1.00
CHR	NC_014751	M. leachii	Mollicutes	1.00
CHR	NC_017521	M. leachii	Mollicutes	1.00
CHR	NC_005364	M. mycoides	Mollicutes	1.00
CHR	NC_015431	M. mucoides	Mollicutes	1.00
CHB	NC 015946	M nutrefaciens	Mollicutes	1.00
	1			
clustor	• 86 stability	measure : 0.991667		
CUD	NC 000000	D other	Dehalterereritet	0.05
CHR	NG_002936	D. etnenogenes	Denalococcoidetes	0.95
CHR	NC_009455	Dehalococcoides sp.	Denalococcoidetes	1.00
CHR	NC_007356	Dehalococcoides sp.	Dehalococcoidetes	1.00
CHR	NC_013552	Dehalococcoides sp.	Dehalococcoidetes	1.00
CHR	NC_013890	Dehalococcoides sp.	Dehalococcoidetes	1.00
	•			
cluster	: 87 stability	measure : 0.672639		
CHR	NC_007677	S. ruber	Unidentified	1.00
CHB	NC 014032	S ruber	Unidentified	1.00
clustor	• 88 stability	measure : 0.770381		
DI	NG 01cc01	D II CONTROLL		1.00
PL	NC_016601	P. aioxanivorans	Actinobacteridae	1.00
PL	NC_004719	S. avermitilis	Actinobacteridae	1.00
PL	NC_003903	S. coelicolor	Actinobacteridae	1.00
\mathbf{PL}	NC_016110	S. flavogriseus	Actinobacteridae	1.00
\mathbf{PL}	NC_016115	S. flavogriseus	Actinobacteridae	1.00
\mathbf{PL}	NC_016972	S. hygroscopicus	Actinobacteridae	0.43
\mathbf{PL}	NC_017766	S. hyproscopicus	Actinobacteridae	0.73
PL.	NC 004933	S lividans	Actinobacteridae	1.00
PI	NC 004808	S rochei	Actinobacteridae	1.00
DI	NC 004024	S. violaccomubor	Actinobactoridae	1.00
	NC 015051	S. Undracedrader	Actinobacteridae	1.00
PL	NC-015951	S. violaceusniger	Actinobacteridae	1.00
PL	NC_010311	Streptomyces sp.	Actinobacteridae	0.73
PL	NC_010851	Streptomyces sp.	Actinobacteridae	0.57
PL	NZ_CM00116	6 Streptomyces sp.	Actinobacteridae	1.00
PL	NC_015409	V. maris	Actinobacteridae	1.00
cluster	•:89 stability	measure : 1.000000		
PL	NC_004308	B. grahamii	α proteobacteria	1.00
\mathbf{PL}	NC_006374	B. grahamii	α proteobacteria	1.00
\mathbf{PL}	NC_010615	B. tribocorum	αproteobacteria	1.00
\mathbf{PL}	NC 010021	Z. mobilis	αproteobacteria	1.00
PL.	NC 017185	Z mobilis	oproteobacteria	1.00
PI	NC 004969	Preudoalteromonas sp	oproteobacteria	1.00
ъ	NC 007197	S alossimidias	protechacteria	1.00
	NC 007189	S. glossiniaius	· yproteobacteria	1.00
PL DI	NC 007715	S. giossiniaius	proteobacteria	1.00
PL	INC_007715	5. giossiniaius	γproteopacteria	nan
		monoune + 0 522057		
ciuster	. 90 stability	measure : 0.533057		0.00
PL	NC_011367	G. diazotrophicus	α proteobacteria	0.36
PL	NC_002033	N. aromaticivorans	α proteobacteria	0.93
\mathbf{PL}	NC_009426	N. aromaticivorans	α proteobacteria	0.94
\mathbf{PL}	NC_015579	Novosphingobium sp.	α proteobacteria	0.88
\mathbf{PL}	NC_008688	P. denitrificans	α proteobacteria	0.23
\mathbf{PL}	NC_011143	P. zucineum	α proteobacteria	0.40
\mathbf{PL}	NC_015595	S. chlorophenolicum	α proteobacteria	0.64
PL	NC 014007	S. japonicum	aproteobacteria	0.91
PL	NC 000507	S wittichii	oproteobactoria	0.77
DI	NC 000509	S wittichii	aproteobactoria	0.86
	NC 015074	Smbinaching	aprotectuation	0.00
	NC-015974	Sphingoorum sp.	aproteobacteria	0.80
PL	NG 018308	Sphingooium sp.	αproteopacteria	0.73
PL	NC-013358	2. mobilis	α proteobacteria	0.50
-				
cluster	•: 91 stability	measure : 1.000000		
\mathbf{PL}	NC_008597	E. coli	γ proteobacteria	1.00
\mathbf{PL}	NC_009789	E. coli	γ proteobacteria	1.00
\mathbf{PL}	NC_011602	E. coli	γ proteobacteria	1.00
				1 00
PL	NC_017658	E. coli	γ proteobacteria	1.00
PL PL	NC_017658 NC_019060	E. coli E. coli	γ proteobacteria γ proteobacteria	1.00
PL PL PL	NC_017658 NC_019060 NC_019070	E. coli E. coli E. coli	γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00
PL PL PL DI	NC_017658 NC_019060 NC_019070	E. coli E. coli E. coli S. enterica	γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00
PL PL PL PL	NC_017658 NC_019060 NC_019070 NC_005862	E. coli E. coli E. coli S. enterica	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00
PL PL PL PL PL	NC_017658 NC_019060 NC_019070 NC_005862 NC_011079	E. coli E. coli E. coli S. enterica S. enterica	γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	1.00 1.00 1.00 1.00 1.00

1 11	NC_017320	S. flexneri	γ proteobacteria	1.00	\mathbf{PL}	NC_004930	L. lactis	Bacilli	1.00
\mathbf{PL}	NC_017329	S. flexneri	γ proteobacteria	1.00	\mathbf{PL}	NC_004981	L. lactis	Bacilli	1.00
					\mathbf{PL}	NC_011610	L. lactis	Bacilli	1.00
cluster	: 92 stability	measure : 1.000000			\mathbf{PL}	NC_013783	L. lactis	Bacilli	1.00
\mathbf{PL}	NC_002635	B. longum	Actinobacteridae	1.00	PL	NC_006399	L. plantarum	Bacilli	1.00
PL	NC_004252	B. longum	Actinobacteridae	1.00	PL	NC_010098	L. plantarum	Bacilli	1.00
PL	NC_004443	B. longum	Actinobacteridae	1.00	PL	NC_012628	L. plantarum	Bacilli Da ailli	1.00
PL	NC_004769	B. longum	Actinobacteridae	1.00		NC 014026	L. plantaram	Bacilli	1.00
PL PL	NC 006843	B. longum	Actinobacteridae	1.00	PL PL	NC 013969	L. reuleri S aureus	Bacilli	1 00
PL	NC 006997	B longum	Actinobacteridae	1.00	PL	NC 013345	S aureus	Bacilli	1.00
PL	NC_010861	B. longum	Actinobacteridae	1.00	PL	NC_005908	S. aureus	Bacilli	1.00
\mathbf{PL}	NC_015066	B. longum	Actinobacteridae	1.00	\mathbf{PL}	NC_001382	M. mycoides	Mollicutes	1.00
\mathbf{PL}	NC_017220	B. longum	Actinobacteridae	1.00	\mathbf{PL}	NC_001843	H. pylori	ϵ proteobacteria	1.00
					\mathbf{PL}	NC_014162	H. pylori	ϵ proteobacteria	1.00
cluster	: 93 stability	measure : 0.900000			\mathbf{PL}	NC_014163	H. pylori	ϵ proteobacteria	1.00
\mathbf{PL}	NC_001910	B. aphidicola	γ proteobacteria	1.00	<u> </u>				
\mathbf{PL}	NC_001911	B. aphidicola	γ proteobacteria	1.00	cluster	: 99 stability 1	measure : 0.487323		
PL	NC_002253	B. aphidicola	γ proteobacteria	1.00	CHR	NC_012108	D. autotrophicum	δ proteobacteria	nan
PL	NC_004555	B. aphidicola	γ proteobacteria	1.00	CHR	NC_018645	D. toluolica	∂proteobacteria	1.00
PL	NC_012540	B. aphiaicola B. anhidicola	γ proteobacteria	1.00	alustan	100 stability	monguro + 0 728111		
PL	NC 017261	B. aphilicola	oproteobacteria	1.00	CHB	NC 013410	F succinogenes	Fibrobacteria	1.00
PL	NC 017257	B. aphilicola	vproteobacteria	1.00	CHR	NC 017448	F succinogenes	Fibrobacteria	1.00
PL	NZ_CM00095	7 Cand. Regiella	γproteobacteria	1.00	01110	1.0101110	1 / oucomogeneo	Throbactoria	1 1.00
\mathbf{PL}	NC_009829	S. proteamaculans	γ proteobacteria	0.33	cluster	: 101 stability	measure : 0.885556		
\mathbf{PL}	NC_007713	S. glossinidius	γ proteobacteria	1.00	CHR	NC_014758	C. nitroreducens	Deferribacteres	1.00
\mathbf{PL}	NC_007182	S. glossinidius	γ proteobacteria	1.00	CHR	NC_013939	D. desulfuricans	Deferribacteres	1.00
PL	NC_007183	S. glossinidius	γ proteobacteria	1.00	\mathbf{CHR}	NC_015672	F. sinusarabici	Deferribacteres	1.00
cluster	: 94 stability	measure : 0.597475			cluster	: 102 stability	measure : 0.486589		
PL	NC_014819	A. excentricus	α proteobacteria	0.60	PL	NC_004954	Micrococcus sp.	Actinobacteridae	0.97
PL DT	NC 011004	A. hpoferum	αproteobacteria	0.45	PL	NC_017081	P. mikurensis	Phycisphaerae	0.00
гь рт	NC 006277	A. tumefaciene	aproteobacteria	1.00	PL DI	NC 015001	Nitrosomanas	pproteobacteria	0.41
PL	NC 013860	A cospirillum SD	aproteobacteria	0.41	PL	NC 015223	Nitrosomonas sp.	Bproteobacteria	0.13
PL	NC_010335	Caulobacter sp.	aproteobacteria	0.41	PL	NC 002759	P suringae	oproteobacteria	1.00
\mathbf{PL}	NC_010727	M. populi	α proteobacteria	0.55	PL	NC_005205	P. syringae	γ proteobacteria	0.88
\mathbf{PL}	NC_011893	M. nodulans	α proteobacteria	0.88	\mathbf{PL}	NC_005918	P. syringae	γ proteobacteria	0.73
\mathbf{PL}	NC_010502	M. radiotolerans	α proteobacteria	nan	\mathbf{PL}	NC_005919	P. syringae	γ proteobacteria	1.00
PL	NC_010504	M. radiotolerans	α proteobacteria	0.92	\mathbf{PL}	NC_007274	P. syringae	$\gamma { m proteobacteria}$	0.57
PL	NC_012810	M. extorquens	α proteobacteria	0.86	\mathbf{PL}	NC_007275	P. syringae	γ proteobacteria	1.00
PL	NC_015582	Novosphingobium sp.	α proteobacteria	0.73	PL	NC_013178	V. alginolyticus	γ proteobacteria	1.00
PL	NC_010124	G. aiazotrophicus	αproteobacteria	0.94	PL	NC_009351	V. anguillarum	γ proteobacteria	1.00
PL PL	NC 008388	r. aminophilas B. depitrificans	aproteobacteria	1.00	PL	NC 003021	V. paranaemolyticus X. azonopodis	γproteobacteria	1.00
PL	NC 016813	S fredii	aproteobacteria	0.75	FL	NC_003921	A. axonopours	yproteobacteria	0.00
PL	NC_016836	S. fredii	αproteobacteria	0.67	cluster	: 103 stability	measure : 0.568096		
DI	NC 013357	Z mobilis	oproteobacteria	0.82	CHR	NC_014484	S. thermophila	Spirochaetes	1.00
PL	110-010001	1. 11000000	aprocobaccoria				······································	· · · · · · · · · · · · · · · · · · ·	
PL PL	NC_013787	Z. mobilis	α proteobacteria	0.79	CHR	NC_017583	S. thermophila	Spirochaetes	1.00
PL PL PL	NC_013787 NC_017181	Z. mobilis Z. mobilis	α proteobacteria α proteobacteria	0.79 0.92	CHR	NC_017583	S. thermophila	Spirochaetes	1.00
PL PL PL PL	NC_013787 NC_017181 NC_010000	Z. mobilis Z. mobilis S. baltica	lphaproteobacteria lphaproteobacteria γ proteobacteria	0.79 0.92 1.00	CHR cluster	NC_017583	S. thermophila measure : 0.642879	Spirochaetes	1.00
PL PL PL PL	NC_013387 NC_017181 NC_010000	Z. mobilis Z. mobilis Z. mobilis S. baltica	α proteobacteria α proteobacteria γ proteobacteria	0.79 0.92 1.00	CHR cluster CHR	NC_017583 : 104 stability NC_011761	S. thermophila measure : 0.642879 A. ferrooxidans	Spirochaetes γ proteobacteria	1.00
PL PL PL PL cluster	NC_013787 NC_017181 NC_010000 : 95 stability	Z. mobilis Z. mobilis S. baltica measure : 0.850350	αproteobacteria αproteobacteria γproteobacteria	0.79 0.92 1.00	CHR cluster CHR CHR	NC_017583 • : 104 stability NC_011761 NC_011206	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans	Spirochaetes γ proteobacteria γ proteobacteria	1.00 1.00
PL PL PL PL cluster PL PL	NC_013387 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve	αproteobacteria αproteobacteria γproteobacteria Actinobacteridae Actinobacteridae	0.79 0.92 1.00	CHR cluster CHR CHR	NC_017583 • : 104 stability NC_011761 NC_011206 • : 105 stability	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans	Spirochaetes γ proteobacteria γ proteobacteria	1.00 1.00
PL PL PL PL cluster PL PL PL	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum	αproteobacteria αproteobacteria γproteobacteria Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43	CHR cluster CHR CHR cluster PL	NC_017583 • : 104 stability NC_011761 NC_011206 • : 105 stability NC_016592	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp	Spirochaetes γproteobacteria γproteobacteria	1.00 1.00
PL PL PL PL PL PL PL PL PL	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010877 NC_013448	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis	Aproteobacteria aproteobacteria yproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae	$ \begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array} $	CHR cluster CHR CHR cHR PL PL	NC_017583 • : 104 stability NC_011761 NC_011206 • : 105 stability NC_016592 NC_010115	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii	Spirochaetes γproteobacteria γproteobacteria βproteobacteria γproteobacteria	1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp.	Aproteobacteria aproteobacteria yproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79 \\ 0.92 \\ 1.00 \\ \hline \\ 0.73 \\ 0.43 \\ 0.43 \\ 1.00 \\ 1.00 \\ \end{array}$	CHR cluster CHR CHR CHR cluster PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_01015 NC_002118	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii	Spirochaetes γproteobacteria γproteobacteria βproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica	Aproteobacteria αproteobacteria γproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79 \\ 0.92 \\ 1.00 \\ \hline \\ 0.73 \\ 0.43 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ \end{array}$	CHR cluster CHR CHR cluster PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_06592 NC_010115 NC_002118 NC_002131	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii	Spirochaetes γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913 NC_010065	2. mobilis 2. mobilis 2. mobilis 3. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii	Aproteobacteria αproteobacteria γproteobacteria γproteobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR cluster PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_00115 NC_002118 NC_002131 NC_004704	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii	Spirochaetes γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913 NC_010065 NC_004526	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR cluster PL PL PL PL PL PL PL	NC_017583 .: 104 stability NC_011761 NC_011206 .: 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii	Spirochaetes γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010937 NC_013448 NC_010877 NC_013448 NC_010657 NC_006258 NC_006258	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_010258	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_010874 NC_013448 NC_010874 NC_004526 NC_004526 NC_006258 NC_008792	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. ghanaensis	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_006592 NC_010115 NC_002118 NC_002131 NC_009726 NC_010258 NC_011526	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913 NC_014913 NC_014913 NC_004526 NC_004526 NC_006258 NC_006258 NC_008792 NC_007431	2. mobilis 2. mobilis 2. mobilis 2. mobilis 5. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_006592 NC_01015 NC_002118 NC_002131 NC_002131 NC_004704 NC_009726 NC_010258 NC_010258 NC_010526 NC_006365	S. thermophila M. ferrooxidans A. ferrooxidans M. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. angezelzi	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010877 NC_014913 NC_010065 NC_006258 NC_006258 NC_006258 NC_006258 NC_0062792 NC_002149 NC_002149 NC_002431	2. mobilis 2. mobilis 2. mobilis 3. baltica measure : 0.850350 B. breve B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomuces sp.	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_002131 NC_004704 NC_000726 NC_010258 NC_011526 NC_006365 NC_006365 NC_006325	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. balirca	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.25 0.45
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010877 NC_014913 NC_0106258 NC_006258 NC_006258 NC_006258 NC_006258 NC_006258 NC_0021491 NC_001787	2. mobilis 2. mobilis 2. mobilis 3. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_00218 NC_002118 NC_002118 NC_004704 NC_009726 NC_010258 NC_010258 NC_010526 NC_006365 NC_008738 NC_009035	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.25 0.43
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_002133 NC_010930 NC_010930 NC_010937 NC_013448 NC_010877 NC_014943 NC_01065 NC_006258 NC_006258 NC_006258 NC_006258 NC_002149 NC_002149 NC_0021431 NC_002187	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002131 NC_004704 NC_009726 NC_004704 NC_009726 NC_004704 NC_009726 NC_006365 NC_008738 NC_009035 : 106 stability	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.25 0.43
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_010877 NC_013448 NC_010877 NC_004526 NC_006258 NC_006258 NC_006258 NC_006258 NC_006258 NC_006258 NC_002149 NC_007431 NC_001787 : 96 stability	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_01526 NC_008738 NC_008738 NC_009035 : 106 stability NC_003091	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.25 0.43
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_010000 : 95 stability NC_002133 NC_010930 NC_010874 NC_013448 NC_010874 NC_013448 NC_013448 NC_013448 NC_014549 NC_007431 NC_001787 : 96 stability	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_009726 NC_010258 NC_010258 NC_010258 NC_008738 NC_009035 : 106 stability NC_003091 NC_017773	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.25 0.43 0.86 0.73
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_010874 NC_014913 NC_004526 NC_004526 NC_004526 NC_004526 NC_004531 NC_001787 : 96 stability NC_004549 NC_008712	Z. mobilis Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_010258 NC_010258 NC_009035 : 106 stability NC_003091 NC_017773	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.25 0.43 0.86 0.73
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_010874 NC_004931 NC_002149 NC_002187 : 96 stability NC_014549 NC_008712 NC_011881 NC_001287	2. mobilis 2. mobilis 2. mobilis 3. mobilis 5. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus A theoret.	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}\\ \hline 0.73\\ 0.43\\ 1.00\\ $	CHR cluster CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000	Spirochaetes γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.25 0.43 0.86 0.73
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_017181 NC_002133 NC_010900 : 95 stability NC_002133 NC_010877 NC_013448 NC_010877 NC_014913 NC_010065 NC_006258 NC_006258 NC_002149 NC_002149 NC_002149 NC_002149 NC_002131 NC_014549 NC_00452	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurileitensis A. chlorophenolicus Arthrobacter sp.	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_009726 NC_004704 NC_009726 NC_006365 NC_006365 NC_006365 NC_006365 NC_008738 NC_009035 : 106 stability NC_003091 NC_017773 : 107 stability	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae	Spirochaetes γproteobacteria Mollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.43 0.86 0.73 1.000
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010877 NC_01065 NC_00655 NC_00655 NC_002149 NC_002149 NC_002149 NC_002149 NC_002149 NC_002149 NC_002131 NC_014549 NC_008537 NC_009453 NC_010494	 Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyccs sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. 	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_001526 NC_006365 NC_008738 NC_009035 : 106 stability NC_012806 NC_003640 NC_002925	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae	Spirochaetes γproteobacteria Nproteobacteria Mollicutes Mollicutes Mollicutes	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010874 NC_013448 NC_013448 NC_013448 NC_014512 NC_004526 NC_004526 NC_004531 NC_004531 NC_014549 NC_008537 NC_009453 NC_014549	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp.	Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_011526 NC_009738 NC_009035 : 106 stability NC_003091 NC_017773 : 107 stability NC_012806 NC_00360 NC_007295 NC_007322	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. byopneumoniae	Spirochaetes γproteobacteria Nolocateria γproteobacteria Nollicutes Mollicutes Mollicutes Mollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.43 0.86 0.73 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010874 NC_013448 NC_01874 NC_014548 NC_004526 NC_004526 NC_004526 NC_004526 NC_004528 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_0074549 NC_007453 NC_008537 NC_009453 NC_010494 NC_018532 NC_010399	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis	Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}\\ \hline 0.73\\ 0.43\\ 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 0.12\\ 0.12\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002131 NC_002131 NC_009726 NC_010258 NC_010258 NC_009035 : 106 stability NC_003091 NC_017773 : 107 stability NC_012806 NC_00360 NC_007392 NC_017509	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquacolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria Mollicutes Mollicutes Mollicutes Mollicutes Mollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.43 0.86 0.73 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_014913 NC_004526 NC_004526 NC_004526 NC_004526 NC_004526 NC_004528 NC_007431 NC_007432 NC_007433NC_007433 NC_007433 NC_00743	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis	Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}\\ \hline 0.73\\ 0.43\\ 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 0.25\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 0.12\\ 0.12\\ 0.83\\ \end{array}$	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_010258 NC_010258 NC_009035 : 106 stability NC_003091 NC_017773 : 107 stability NC_003091 NC_017295 NC_007332 NC_007332 NC_017509	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria Nproteobacteria Nproteobacteria Nproteobacteria Nproteobacteria Nproteobacteria Nproteobacteria Nproteobacteria Nproteobacteria Nproteobacteria Nollicutes Mollicutes Mollicutes Mollicutes Mollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.25 0.25 0.43 0.86 0.73 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_014913 NC_004526 NC_004526 NC_004526 NC_004526 NC_004528 NC_004531 NC_004931 NC_014549 NC_014549 NC_014549 NC_008517 NC_009453 NC_010494 NC_018532 NC_010494 NC_010499 NC_010408 NC_010408 NC_010408	Z. mobilis Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. C. michiganensis C. michiganensis R. jostii	Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}\\ \hline 0.73\\ 0.43\\ 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 0.12\\ 0.83\\ 0.75\\ \end{array}$	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL PL	NC_017583 :: 104 stability NC_011761 NC_011206 :: 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_010258 NC_010526 NC_006365 NC_006365 NC_006365 NC_006365 NC_007332 NC_017509 : 108 stability	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria %proteobacteria %proteobacteria </th <th>1.00 1.00</th>	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_00930 NC_010930 NC_010877 NC_013448 NC_010877 NC_014913 NC_010065 NC_00655 NC_00655 NC_002149 NC_002149 NC_002149 NC_002149 NC_0021431 NC_014549 NC_004712 NC_011881 NC_008537 NC_009453 NC_010494 NC_018532 NC_010494 NC_008270 NC_009479	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis	Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}\\ \hline 0.73\\ 0.43\\ 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 0.12\\ 0.12\\ 0.12\\ 0.83\\ 0.75\\ 0.19\\ \end{array}$	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_001526 NC_006365 NC_006365 NC_006365 NC_006365 NC_003091 NC_017773 : 107 stability NC_012806 NC_007332 NC_007332 NC_017509 : 108 stability NC_007101	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria Mollicutes Mollicutes Mollicutes Mollicutes Mollicutes Mollicutes	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913 NC_01065 NC_004526 NC_00655 NC_004526 NC_002149 NC_007431 NC_002149 NC_007431 NC_002149 NC_002149 NC_002431 NC_014549 NC_008537 NC_014549 NC_008537 NC_009453 NC_010494 NC_018532 NC_010498 NC_009479 NC_009479 NC_002520	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. opacus	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002131 NC_004704 NC_009726 NC_001526 NC_004704 NC_009035 : 106 stability NC_003091 NC_017773 : 107 stability NC_012806 NC_0036360 NC_007295 NC_007332 NC_017509 : 108 stability NC_007101 NC_007387	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria %proteobacteria %ollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.43 0.86 0.73 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010874 NC_013448 NC_010874 NC_013448 NC_014913 NC_014526 NC_004526 NC_004526 NC_004528 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_004537 NC_004537 NC_009453 NC_010494 NC_018532 NC_010498 NC_010498 NC_008270 NC_009479 NC_002520 NC_003531	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica	Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}\\ \hline 0.73\\ 0.43\\ 0.43\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.25\\ 1.00\\ 1.00\\ 0.86\\ 0.62\\ 0.12\\ 0.12\\ 0.83\\ 0.75\\ 0.19\\ man\\ 0.80\\ cc\\ cc\\ cc\\ cc\\ cc\\ cc\\ cc\\ cc\\ cc\\ c$	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_010258 NC_009726 NC_009738 NC_009035 : 106 stability NC_012806 NC_003091 NC_017773 : 107 stability NC_012806 NC_007322 NC_007322 NC_00737 NC_007387 NC_007387 NC_007388	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria Mollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.86 0.73 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_014913 NC_004526 NC_004526 NC_004526 NC_004526 NC_004528 NC_004531 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_004537 NC_009453 NC_010494 NC_018532 NC_010499 NC_010408 NC_008270 NC_009479 NC_012520 NC_013531 NC_014304	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae	Actinobacteridae Actinobacteridae	$\begin{array}{c} 0.79\\ 0.92\\ 1.00\\ \end{array}\\ \hline 0.73\\ 0.43\\ 0.43\\ 1.00\\ 0.86\\ 0.62\\ 0.12\\ 0.12\\ 0.12\\ 0.00\\ $	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002131 NC_002131 NC_009726 NC_009726 NC_010258 NC_010258 NC_009035 : 106 stability NC_003091 NC_017773 : 107 stability NC_007382 NC_007387 NC_007388 NC_007389	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria Nollicutes Mollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.86 0.73 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_014913 NC_004526 NC_004526 NC_004526 NC_004526 NC_004526 NC_004531 NC_004531 NC_014549 NC_004537 NC_009453 NC_010494 NC_018532 NC_010494 NC_018532 NC_010494 NC_008537 NC_009453 NC_010494 NC_008537 NC_009453 NC_010494 NC_008537 NC_009453 NC_010494 NC_008537 NC_009453 NC_010494 NC_008537 NC_009453 NC_010494 NC_018532 NC_010494 NC_008537 NC_009453 NC_010494 NC_008537 NC_009453 NC_010494 NC_008537 NC_009453 NC_010494 NC_012520 NC_013531 NC_014304	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002131 NC_004704 NC_009726 NC_010258 NC_010258 NC_006365 NC_006365 NC_006365 NC_006365 NC_007035 : 106 stability NC_012806 NC_007392 NC_007332 NC_007322 NC_007387 NC_007387 NC_007388 NC_007389 NC_007390	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria %proteobacteria %ollicutes Mollicutes Mollicutes Mollicutes <t< th=""><th>1.00 1.00</th></t<>	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_00930 NC_010930 NC_010877 NC_013448 NC_010877 NC_010877 NC_00877 NC_0004526 NC_00655 NC_004526 NC_002149 NC_002149 NC_002149 NC_002149 NC_002431 NC_004931 NC_014549 NC_004537 NC_014549 NC_004537 NC_014549 NC_004537 NC_014549 NC_018532 NC_010494 NC_018532 NC_010494 NC_018532 NC_010494 NC_008270 NC_002520 NC_014304 : 97 stability	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae measure : 0.412099	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002118 NC_002131 NC_004704 NC_009726 NC_00726 NC_004704 NC_00365 NC_006365 NC_006365 NC_006365 NC_006360 NC_0073091 NC_012806 NC_007392 NC_007387 NC_007387 NC_007390 NC_007391	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria Nollicutes Mollicutes	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913 NC_01065 NC_004526 NC_006258 NC_006258 NC_006258 NC_002149 NC_002149 NC_002149 NC_002149 NC_002431 NC_014549 NC_008537 NC_014549 NC_008537 NC_014549 NC_008537 NC_014549 NC_008537 NC_014532 NC_014532 NC_01408 NC_009479 NC_002520 NC_009479 NC_012520 NC_013531 NC_014304 : 97 stability NC_016884	 Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. preve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae measure : 0.412099 S. acidophilus S. acidophilus 	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002118 NC_002101 NC_004704 NC_009726 NC_001526 NC_006365 NC_006365 NC_006365 NC_006365 NC_006365 NC_007381 : 107 stability NC_012806 NC_007322 NC_007322 NC_007387 NC_007387 NC_007381 NC_007391 . 100 stability	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria Nollicutes Mollicutes	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_010874 NC_014913 NC_014526 NC_004526 NC_004526 NC_004526 NC_004531 NC_007431 NC_007431 NC_004931 NC_007431 NC_004931 NC_007431 NC_004931 NC_007431 NC_004537 NC_009453 NC_010454 NC_018532 NC_010399 NC_010408 NC_009479 NC_012520 NC_009479 NC_012520 NC_013531 NC_014304 : 97 stability NC_015757 NC_016884	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae measure : 0.412099 S. acidophilus	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.0	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002131 NC_004704 NC_009726 NC_010258 NC_009726 NC_001526 NC_00365 NC_008738 NC_009035 : 106 stability NC_012806 NC_003091 NC_017773 : 107 stability NC_012806 NC_007332 NC_007389 NC_007389 NC_007389 NC_007391 : 109 stability NC_013201	S. thermophila measure : 0.642879 A. ferroozidans A. ferroozidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria %proteobacteria %ollicutes Mollicutes Mollicutes Mollicutes <t< th=""><th>1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.43 0.86 0.73 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00</th></t<>	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.43 0.86 0.73 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_013448 NC_010874 NC_004526 NC_004526 NC_004526 NC_004526 NC_004526 NC_004526 NC_004539 NC_007431 NC_007431 NC_007431 NC_007431 NC_004931 NC_007431 NC_004931 NC_007454 NC_008537 NC_009453 NC_010494 NC_018532 NC_010494 NC_018532 NC_010498 NC_008537 NC_009453 NC_010494 NC_018532 NC_010494 NC_012520 NC_013531 NC_014304 : 97 stability NC_015757 NC_016884 : 98 stability	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae measure : 0.412099 S. acidophilus S. acidophilus measure : 1.000000	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00	CHR cluster CHR CHR CHR PL PL PL PL PL PL PL PL PL CLUSTER CHR CHR CHR CHR CHR CHR CHR CH	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002131 NC_004704 NC_009726 NC_009726 NC_009726 NC_001526 NC_008738 NC_009035 : 106 stability NC_003091 NC_017773 : 107 stability NC_007302 NC_007322 NC_007388 NC_007388 NC_007388 NC_007389 NC_007390 NC_007391 : 109 stability NC_015966	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. citri S. citri	Spirochaetes γproteobacteria Mollicutes	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.25 0.43 0.86 0.73 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_013787 NC_013787 NC_013787 NC_013787 NC_01131 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913 NC_014913 NC_004526 NC_004526 NC_004526 NC_004526 NC_004526 NC_00453 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_007431 NC_004931 NC_007453 NC_009453 NC_010494 NC_018532 NC_010494 NC_012520 NC_012520 NC_012520 NC_012557 NC_016884 : 98 stability NC_004985	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. natalensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae measure : 0.412099 S. acidophilus S. acidophilus	Actinobacteria Actinobacteria Actinobacteria Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_002118 NC_002118 NC_002131 NC_004704 NC_009726 NC_009726 NC_009035 : 106 stability NC_003091 NC_012806 NC_007395 NC_007392 NC_007322 NC_007322 NC_007325 NC_007388 NC_007389 NC_007389 NC_007390 : 109 stability NC_013501 NC_013501 NC_013501 NC_013501	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria %proteobacteria %ollicutes	1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_013787 NC_013787 NC_013787 NC_017181 NC_010000 : 95 stability NC_002133 NC_010930 NC_010877 NC_013448 NC_010874 NC_014913 NC_014913 NC_004526 NC_004526 NC_004526 NC_004526 NC_004526 NC_004528 NC_004931 NC_0014549 NC_007431 NC_004931 NC_004931 NC_004931 NC_014549 NC_007431 NC_004931 NC_014549 NC_014549 NC_014549 NC_014549 NC_014532 NC_010494 NC_010494 NC_010494 NC_012520 NC_014304 : 97 stability NC_015757 NC_016884 : 98 stability	2. mobilis 2. mobilis 2. mobilis 3. baltica measure : 0.850350 B. breve B. pseudolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. natalensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. S. michiganensis R. jostii C. michiganensis R. jostii C. michiganensis R. jostii C. michiganensis R. jostii C. michiganensis S. acidophilus S. acidophilus S. acidophilus L. buchneri	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002118 NC_004704 NC_009726 NC_010258 NC_010258 NC_006365 NC_006365 NC_006365 NC_006365 NC_007391 : 107 stability NC_012806 NC_007392 NC_007322 NC_007322 NC_007322 NC_007387 NC_007387 NC_007387 NC_007389 NC_007390 : 109 stability NC_015966 : 110 stability	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii L. pneumophila M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria %proteobacteria %ollicutes Mollicutes Mollicutes	1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_013787 NC_013787 NC_013787 NC_017181 NC_002133 NC_010930 NC_010930 NC_010877 NC_013448 NC_010877 NC_014913 NC_01065 NC_00655 NC_006258 NC_006258 NC_002149 NC_002149 NC_002149 NC_002149 NC_002149 NC_002149 NC_002431 NC_014549 NC_004931 NC_0104531 NC_010494 NC_018532 NC_010494 NC_008270 NC_009479 NC_008270 NC_009479 NC_008270 NC_002520 NC_014304 : 97 stability NC_014304 : 97 stability NC_016884 : 98 stability	Z. mobilis Z. mobilis Z. mobilis S. baltica measure : 0.850350 B. breve B. psevelolongum N. aobensis Nocardia sp. P. autotrophica P. freudenreichii P. granulosum R. erythropolis S. ghanaensis S. venezuelae Streptomyces sp. T. pyogenes measure : 0.413470 A. arilaitensis A. aurescens A. chlorophenolicus Arthrobacter sp. Arthrobacter sp. C. michiganensis R. jostii C. michiganensis R. jostii C. michiganensis R. opacus X. cellulosilytica E. billingiae measure : 0.412099 S. acidophilus S. acidophilus L. buchneri L. casei	Actinobacteridae Actinobacteridae	0.79 0.92 1.00 0.73 0.43 0.43 1.00	CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017583 : 104 stability NC_011761 NC_011206 : 105 stability NC_016592 NC_010115 NC_002118 NC_002118 NC_002104 NC_009726 NC_010258 NC_010258 NC_006365 NC_006365 NC_006365 NC_006365 NC_003091 NC_017773 : 107 stability NC_012806 NC_007302 NC_007322 NC_017509 : 108 stability NC_007387 NC_007387 NC_007389 NC_007391 : 109 stability NC_012506 : 110 stability NC_014008	S. thermophila measure : 0.642879 A. ferrooxidans A. ferrooxidans measure : 0.610106 Burkholderia sp. C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii C. burnetii M. aquaeolei S. baltica measure : 0.312042 F. nucleatum R. aquatilis measure : 1.000000 M. conjunctivae M. hyopneumoniae M. hyopneumoniae	Spirochaetes γproteobacteria %proteobacteria %ollicutes Mollicutes Mollicutes Mollicutes Mollicutes	1.00 1.00

CHR	NC_010571	O. terrae	Opitutae	0.54	\mathbf{PL}	NC_010643	P. rettgeri
	•				\mathbf{PL}	NC_009982	S. enterica
cluster :	111 stability	7 measure : 0.319796			\mathbf{PL}	NC_010716	S. enterica
PL	NC_006909	L. ferrooxidans	Nitrospira	0.50	. .		
PL	NC_015592	S. meliloti	α proteobacteria	1.00	cluster	: 122 stability	measure : (
PL	NC 000228	Z. moonis P. vietnamieneia	aproteobacteria	0.45	DI	NC_009972	H. auranti
PL	NC 008761	P. naphthalenivorans	βproteobacteria	0.60	PL	NC 009974	H. auranti
PL	NC_008013	L. intracellularis	δ proteobacteria	0.56			
\mathbf{PL}	NC_010600	A. caldus	γ proteobacteria	0.50	cluster	: 123 stability	measure : (
\mathbf{PL}	NC_015854	A. caldus	γ proteobacteria	0.86	PL	NC_009932	A. marina
\mathbf{PL}	NC_013862	A. vinosum	γ proteobacteria	0.33	\mathbf{PL}	NC_010631	N. punctife
\mathbf{PL}	NC_019092	E. coli	γ proteobacteria	0.71	\mathbf{PL}	NC_010542	Cyanothecological contract of the contract o
PL	NC_009704	Y. pseudotuberculosis	γ proteobacteria	0.71	\mathbf{PL}	NC_013168	Cyanothecological contract of the contract o
PL	NC_006385	bacterium	Unidentified	nan	PL	NC_005232	Synechocys
almatan	110 -t - Lilit.				PL	NC_005229	Synechocys
DI	NC 005244	P metida	erprotochastoria	1.00	PL	NC_008440	D achiera
PL PI	NC 008275	r. putida	γproteobacteria	1.00	PL	NC 015180	D. goorens
PL	NC 011838	P nutida	vproteobacteria	1.00	11	101010100	A. manuo
PL	NC_014124	P. putida	γproteobacteria	1.00	cluster	: 124 stability	measure : (
\mathbf{PL}	NC_015855	P. putida	γ proteobacteria	1.00	PL	NC_001425	S. flavovir
\mathbf{PL}	NC_018746	P. putida	γ proteobacteria	1.00	\mathbf{PL}	NC_010097	S. lavendu
\mathbf{PL}	NC_004444	P. resinovorans	γ proteobacteria	0.50	\mathbf{PL}	NC_008441	S. laurenti
\mathbf{PL}	NC_004633	P. syringae	γ proteobacteria	0.79	\mathbf{PL}	NC_013596	$S. \ roseum$
• .					PL	NC_001759	S. phaeoch
cluster :	113 stability	7 measure : 1.000000	<u> </u>	1 0 0	PL	NC_010849	Streptomy
CHR	NC_011205	C. exile	Caldisericia	1.00	PL	NC_013417	Streptomy
CIII	NC_011295	C. proteolyticus	Clostifula	1.00	PL	NC 013667	Streptomy
cluster :	114 stability	7 measure : 0.574386			112	1101010007	Streptomy
PL	NC_011727	Cuanothece sp.	Chroobacteria	1.00	cluster	: 125 stability	measure : 1
\mathbf{PL}	NC_011880	Cyanothece sp.	Chroobacteria	1.00	CHR	NC_011661	D. turgidu
\mathbf{PL}	NC_014534	Cyanothece sp.	Chroobacteria	1.00	CHR	NC_011297	D. thermo
\mathbf{PL}	NC_007411	A. variabilis	Homogonae	0.27			
\mathbf{PL}	NC_007410	A. variabilis	Homogonae	1.00	cluster	: 126 stability	measure : (
\mathbf{PL}	NC_014249	N. azollae	Homogonae	1.00	\mathbf{PL}	NC_016975	L. casei
PL	NC_010632	N. punctiforme	Homogonae	nan	PL	NC_004947	L. ferment
PL	NC_003276	Nostoc sp.	Homogonae	1.00	PL	NC_002799	L. lactis
alustan	115 stability	mongung + 0.281055			PL	NC_010913	L. paracas
BECE	NC 003063	A fabrum	oproteobacteria	0.67	PL	NC 013544	L. paracas
RECE	NC_015508	Aarobacterium sp.	αproteobacteria	0.95	PL	NC_013789	L. plantar
\mathbf{PL}	NC_014005	S. japonicum	α proteobacteria	0.56	\mathbf{PL}	NC_011223	L. rhamno
						•	
cluster :	116 stability	measure : 0.408729			cluster	: 127 stability	⁷ measure : (
PL	NC_015147	A. phenanthrenivorans	Actinobacteridae	0.23	PL	NC_015561	A. subflavi
PL	NC_011888	M. nodulans	αproteobacteria	0.64	PL	NC_009339	M. gilvum
PL	NC 000226	R. peacockii	aproteobacteria	0.73	PL	NC 008704	Mycobacte
PL	NC 016819	B aquatilis	oproteobacteria	0.30	PL	NC 008269	R iostii
PL	NC_007094	A. porcitonsillarum	γproteobacteria	1.00	1.2	1101000200	10. 50500
\mathbf{PL}	NC_007095	A. porcitonsillarum	γ proteobacteria	1.00	cluster	: 128 stability	measure :
\mathbf{PL}	NC_007096	A. porcitonsillarum	γ proteobacteria	1.00	PL	NC_006257	L. casei
\mathbf{PL}	NC_009623	A. porcitonsillarum	γ proteobacteria	1.00	\mathbf{PL}	NC_004566	L. ferment
PL	NC_009624	A. porcitonsillarum	γ proteobacteria	0.93	\mathbf{PL}	NC_012222	L. paracas
PL	NC_008538	Arthrobacter sp.	Actinobacteridae	0.18	PL	NC_010603	L. reuteri
alustan	117 stability	monouro + 0 218700			PL	NC_015700	L. reuteri
DI	NC 010270	L homehomeeneis	Bratashaataria	0.77		NC_012042	S. parasan
PL	NC 014144	Thiomonas sp	βproteobacteria	0.57	ГL	NC-015870	5. pseudop
PL	NC_015853	A. caldus	γ proteobacteria	0.64	cluster	: 129 stability	measure : (
\mathbf{PL}	NC_015872	E. coli	γ proteobacteria	0.43	PL	NC_018878	B. thuring
\mathbf{PL}	NC_019163	K. pneumoniae	γ proteobacteria	0.58	\mathbf{PL}	NC_012654	C. botulini
\mathbf{PL}	NC_003922	X. axonopodis	γ proteobacteria	0.55	\mathbf{PL}	NC_010379	C. botulini
\mathbf{PL}	NC_010872	X. axonopodis	γ proteobacteria	0.29	\mathbf{PL}	NC_010418	C. botulini
		•					
PL	NC_010876	X. axonopodis	γ proteobacteria	1.00	PL	NC_010680	C. botulini
$_{\rm PL}$	NC_010876 NC_007507	X. axonopodis X. campestris	γ proteobacteria γ proteobacteria	1.00 0.19	$_{\rm PL}^{\rm PL}$	NC_010680 NC_017177	C. botulinı C. difficile
PL PL	NC_010876 NC_007507	X. axonopodis X. campestris	γ proteobacteria γ proteobacteria	1.00 0.19	PL PL	NC_010680 NC_017177	C. botulinu C. difficile
PL PL cluster :	NC_010876 NC_007507 118 stability	X. axonopodis X. campestris measure : 1.000000	γproteobacteria γproteobacteria	1.00 0.19	PL PL <u>cluster</u>	NC_010680 NC_017177	C. botulini C. difficile
PL PL cluster : PL PL	NC_010876 NC_007507 118 stability NC_002114 NC_010403	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii	γ proteobacteria γ proteobacteria β proteobacteria γ proteobacteria	1.00 0.19	PL PL 	NC_010680 NC_017177 • : 130 stability NC_009472 NC_009474	C. botulina C. difficile <u>measure : (</u> A. cryptum A. cryptum
PL PL cluster : PL PL PL PL	NC-010876 NC-007507 118 stability NC-002114 NC-010403 NC-016977	X. axonopodis X. campestris Measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii	γ proteobacteria γ proteobacteria β proteobacteria γ proteobacteria γ proteobacteria	1.00 0.19	PL PL <u>cluster</u> PL PL PL	NC_010680 NC_017177 • : 130 stability NC_009472 NC_009474 NC_008691	C. botulina C. difficile <u>measure : (</u> A. cryptun A. cryptun A. multivo
PL PL Cluster : PL PL PL PL PL	NC-010876 NC-007507 118 stability NC-002114 NC-010403 NC-016977 NC-002175	X. axonopodis X. campestris 7 measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica	γ proteobacteria γ proteobacteria β proteobacteria γ proteobacteria γ proteobacteria	1.00 0.19 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL	NC_010680 NC_017177 NC_009472 NC_009474 NC_008691 NC_015187	C. botulina C. difficile <u>v measure : (</u> A. cryptun A. cryptun A. multivo A. multivo
PL PL Cluster : PL PL PL PL PL PL	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida	γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.19 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL	NC_010680 NC_017177 NC_009472 NC_009474 NC_008691 NC_015187 NC_015189	C. botulina C. difficile <u>measure : (</u> A. cryptum A. cryptum A. multivo A. multivo A. multivo
PL PL Cluster : PL PL PL PL PL PL	NC_010876 NC_007507 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518	X. axonopodis X. campestris 7 measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 0.19\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ $	PL PL Cluster PL PL PL PL PL PL	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015189 NC_006676	C. botulina C. difficile <u>measure : (</u> A. cryptum A. cryptum A. multivo A. multivo G. oxydan.
PL PL Cluster : PL PL PL PL PL Cluster :	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 : 119 stability	X. axonopodis X. campestris 7 measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida 7 measure : 1.000000	γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.19 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL	NC_010680 NC_017177 NC_009472 NC_009474 NC_008691 NC_015187 NC_015187 NC_006676 NC_014009	C. botulinn C. difficile M. cryptun A. cryptun A. multivo A. multivo G. oxydan. S. japonico
PL PL PL PL PL PL PL PL Cluster : PL	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 : 119 stability NC_003090	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.19 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015189 NC_006676 NC_014009 NC_006826	C. botulina C. difficile measure : (A. cryptun A. cryptun A. multivo A. multivo G. oxydan S. japonica S. xenopho
PL PL Cluster : PL PL PL PL PL PL PL PL PL	NC_010876 NC_007507 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 119 stability NC_003090 NC_001385	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum	γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae	1.00 0.19 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_015189 NC_006676 NC_014009 NC_006826 NC_008246	C. botulina C. difficile T measure : (A. cryptum A. cryptum A. multivo A. multivo A. multivo G. oxydan. S. japonica S. xenopho S. yanoiku
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_010403 NC_002175 NC_002175 NC_002175 NC_002518 : 119 stability NC_003090 NC_001385 NC_002115	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum	 γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae 	1.00 0.19 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015189 NC_006676 NC_014009 NC_006826 NC_008246 NC_007505	C. botulina C. difficile (measure : (A. cryptum A. cryptum A. multivo A. multivo A. multivo G. oxydan. S. japonica S. xenopho S. yanoiku X. campesi
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 : 119 stability NC_003090 NC_001385 NC_002115 NC_010243	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum	 γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae 	$\begin{array}{c} 1.00\\ 0.19\\ \hline \\ 1.00\\ 1$	PL PL PL PL PL PL PL PL PL PL PL Clustor	NC_010680 NC_017177 : 130 stability NC_009472 NC_009474 NC_008691 NC_015187 NC_015187 NC_015189 NC_006876 NC_006826 NC_008246 NC_007505 : 131 stability	C. botulina C. difficile (measure : (A. cryptun A. cryptun A. multivo G. axydan S. japonicu S. xenopha S. yanoiku X. campesu
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 : 119 stability NC_003090 NC_001385 NC_002115 NC_010243 NC_004533	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum	 γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae 	$\begin{array}{c} 1.00\\ 0.19\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$	PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010680 NC.017177 : 130 stability NC.009472 NC.009474 NC.008691 NC.015187 NC.015187 NC.015189 NC.006876 NC.006826 NC.008246 NC.008246 NC.007505 : 131 stability NC.007104	C. botulian C. difficile M. cryptun A. cryptun A. cryptun A. multivo A. multivo G. oxydan. S. japonicu S. xenopha S. yanoiku X. campesi M. cereve
PL PL Cluster : PL PL PL PL PL PL PL PL PL PL Cluster :	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 : 119 stability NC_003090 NC_001385 NC_002115 NC_002115 NC_010243 NC_004533 : 120 stability	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum C. glutamicum	 γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae 	$\begin{array}{c} 1.00\\ 0.19\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$	PL PL PL PL PL PL PL PL PL PL Cluster PL PL	NC_010680 NC_017177 : 130 stability NC_009472 NC_009474 NC_008691 NC_015187 NC_015187 NC_006676 NC_014009 NC_006826 NC_007505 : 131 stability NC_007104 NC_016793	C. botulina C. difficile M. cryptum A. cryptum A. cryptum A. multivo A. multivo G. oxydan S. japonica S. yanoiku X. campesi B. cereus B. cereus
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 : 119 stability NC_003090 NC_001385 NC_002115 NC_002115 NC_010243 NC_004533 : 120 stability NC_013523	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum S. dutamicum	 γproteobacteria γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae 	1.00 0.19 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_00676 NC_014009 NC_006826 NC_007505 : 131 stability NC_007104 NC_005704	C. botulina C. difficile M. cryptun A. cryptun A. multivo A. multivo G. oxydan. S. japonica S. xenopha S. yanoiku X. campess measure : B. ccreus B. ccreus B. mycoidd
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002175 NC_002518 : 119 stability NC_001385 NC_001385 NC_004533 : 120 stability NC_013523 NC_011959	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum C. glutamicum M. there i 0.410161 S. thermophilus T. roseum	γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae	$\begin{array}{c} 1.00\\ 0.19\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_006826 NC_014009 NC_006826 NC_008246 NC_007505 : 131 stability NC_007104 NC_005704 NC_004059	C. botulina C. difficile M. cryptum A. cryptum A. cryptum A. multivo A. multivo A. multivo G. oxydan. S. japonica S. yanoiku X. campesi M. cereus B. cereus B. cereus B. mycoid B. thuring
PL PL PL PL PL PL PL PL PL PL PL PL PL CHR CHR	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_010403 NC_002175 NC_002175 NC_002175 NC_002175 NC_001385 NC_001385 NC_001385 NC_002115 NC_004533 : 120 stability NC_013523 NC_011959	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum C. glutamicum M. thalassica S. thermophilus T. roseum	γproteobacteria Actinobacteridae	$\begin{array}{c} 1.00\\ 0.19\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_006826 NC_014009 NC_006826 NC_008246 NC_007505 : 131 stability NC_007104 NC_007104 NC_016793 NC_004059 NC_004059 NC_006821	C. botulina C. difficile A. cryptum A. cryptum A. cryptum A. multivo A. multivo A. multivo G. oxydan. S. japonica S. xenopho S. yanoiku X. campesi B. cereus B. cereus B. cereus B. thuring B. thuring
PL PL PL PL PL PL PL PL PL PL PL PL PL CHR CHR CHR CLUSTER	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_002175 NC_002175 NC_002175 NC_002518 : 119 stability NC_001385 NC_002115 NC_002135 NC_002135 NC_004533 : 120 stability NC_013523 NC_011959 : 121 stability	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum S. thermophilus T. roseum	 γproteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Actinobacteridae 	1.00 0.19 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_015189 NC_006826 NC_008246 NC_008246 NC_007505 : 131 stability NC_007104 NC_016793 NC_004059 NC_006821 NC_019016	C. botulina C. difficile A. cryptum A. cryptum A. cryptum A. multivo G. oxydan. S. japonicu S. xenopho S. yanoiku X. campest M. cereus B. cereus B. cereus B. cereus B. cereus B. thuring F. limi
PL PL PL PL PL PL PL PL PL PL PL PL PL CHR CHR CHR CHR CLuster : PL	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002175 NC_002518 : 119 stability NC_003090 NC_001385 NC_002115 NC_010243 NC_004533 : 120 stability NC_013523 NC_011959 : 121 stability NC_010908	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum S. thermophilus T. roseum measure : 0.750685 B. asteroides	γproteobacteria Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Thermomicrobia Thermomicrobia Actinobacteridae	1.00 0.19 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_009474 NC_008691 NC_015187 NC_015187 NC_006676 NC_014009 NC_006826 NC_007505 : 131 stability NC_007104 NC_016793 NC_005704 NC_004059 NC_006821 NC_019016 NC_010983	C. botulian C. botulian C. difficile A. cryptum A. cryptum A. multivo G. axydan. S. japonicu S. xenopho S. yanoiku X. campest B. cereus B. cereus B. cereus B. cereus B. cereus B. thuring F. limi aproteoba
PL PL PL PL PL PL PL PL PL PL PL PL CHR CHR CHR CHR CHR CHR PL PL PL PL	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002518 : 119 stability NC_003090 NC_001385 NC_001385 NC_001243 NC_004533 : 120 stability NC_013523 NC_011959 : 121 stability NC_010908 NC_004768	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum S. thermophilus T. roseum measure : 0.750685 B. asteroides B. longum P. measure	γproteobacteria Actinobacteridae	1.00 0.19 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_006676 NC_014009 NC_006826 NC_007505 : 131 stability NC_007104 NC_005704 NC_005704 NC_005704 NC_004059 NC_006821 NC_019983	C. botulina C. difficile A. cryptun A. cryptun A. multivo A. multivo G. oxydan. S. japonica S. xenopha S. yanoiku X. campessi Teasure : B. ccreus B. ccreus B. mycoidd B. thuring F. limi aproteoba
PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_016977 NC_002175 NC_002175 NC_002175 NC_001385 NC_001385 NC_001385 NC_004533 : 120 stability NC_013523 NC_013523 NC_013523 NC_01359 : 121 stability NC_010908 NC_004768 NC_004768 NC_001755	X. axonopodis X. campestris Measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida Measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum C. glutamicum C. glutamicum Measure : 0.410161 S. thermophilus T. roseum Measure : 0.750685 B. asteroides B. longum R. marinus P. marinus	γproteobacteria Actinobacteridae Unidentified Unidentified	1.00 0.19 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.00 0.00	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_006826 NC_008246 NC_008246 NC_007505 : 131 stability NC_007104 NC_007104 NC_00704 NC_006821 NC_005704 NC_006821 NC_006821 NC_006821 NC_006823 : 132 stability NC_016082	C. botulina C. difficile M. cryptum A. cryptum A. cryptum A. multivo A. multivo A. multivo G. oxydam. S. japonica S. xenopho S. yanoiku X. campesi B. cereus B. cereus B. cereus B. measure : 1 B. thuring B. thuring B. thuring F. limi aproteoba
PL PL PL PL PL PL PL PL PL PL PL Cluster : CHR CHR CLUSTER : PL PL PL PL PL PL	NC_010876 NC_007507 : 118 stability NC_002114 NC_010403 NC_010403 NC_002175 NC_002175 NC_002175 NC_002175 NC_001385 NC_001385 NC_001385 NC_004533 : 120 stability NC_010533 NC_011959 : 121 stability NC_004768 NC_004768 NC_004755 NC_013502 NC_015970	X. axonopodis X. campestris measure : 1.000000 Nitrosomonas sp. A. baumannii A. baumannii M. thalassica P. putida measure : 1.000000 C. callunae C. glutamicum C. glutamicum C. glutamicum C. glutamicum C. glutamicum S. thermophilus T. roseum measure : 0.750685 B. asteroides B. longum R. marinus R. marinus R. marinus	γproteobacteria γproteobacteridae Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\ 0.19\\ \hline \\ 1.00\\ 1$	PL PL PL PL PL PL PL PL PL PL PL PL PL P	NC_010680 NC_017177 : 130 stability NC_009472 NC_008691 NC_015187 NC_015187 NC_006826 NC_006826 NC_008246 NC_007505 : 131 stability NC_007104 NC_007104 NC_007704 NC_006821 NC_004059 NC_006821 NC_010983 : 132 stability NC_016982 NC_000435	C. botulina C. botulina C. difficile A. cryptur A. cryptur A. cryptur A. multivo A. multivo G. oxydan S. japonicu S. yanoiku X. campes T. measure : B. cereus B. cereus B. cereus B. cereus B. thuring F. limi aproteoba t. garviea L. garviea

	NC_010643	P. rettaeri	γ proteobacteria	0.50
PL	NC_009982	S. enterica	γproteobacteria	1.00
PL	NC_010716	S. enterica	γ proteobacteria	0.77
1				
luster	: 122 stability	measure : 0.713333		
CHR	NC_009972	H. aurantiacus	Chloroflexi	1.00
PL	NC_009973	H. aurantiacus	Chloroflexi	1.00
PL	NC_009974	H. aurantiacus	Chloroflexi	1.00
luster	: 123 stability	measure : 0.548945		
PL	NC_009932	A. marina	Chroobacteria	1.00
PL	NC_010631	N. punctiforme	Chroobacteria	0.90
PL	NC_010542	Cyanothece sp.	Chroobacteria	1.00
PL	NC_013168	Cyanothece sp.	Chroobacteria	1.00
PL	NC_005232	Synechocystis sp.	Chroobacteria	nan
L I	NC_005229	Synechocystis sp.	Chroobacteria	0.63
	NC_008440	Nostoc sp.	Homogonae	0.90
	NC_017793	D. gobiensis	Deinococci	0.50
L	NC_015180	A. multivorum	α proteobacteria	0.60
lustor	• 124 stability	measure : 0.720733		
	NC 001425	S flavovirens	Actinobacteridae	1.00
PL	NC_010097	S. lavendulae	Actinobacteridae	1.00
PL	NC_008441	S. laurentii	Actinobacteridae	1.00
PL	NC_013596	S. roseum	Actinobacteridae	1.00
PL	NC_001759	S. phaeochromogenes	Actinobacteridae	1.00
PL	NC_010849	Streptomyces sp.	Actinobacteridae	1.00
PL	NC_013417	Streptomyces sp.	Actinobacteridae	1.00
PL	NC_013449	Streptomyces sp.	Actinobacteridae	0.33
PL	NC_013667	Streptomyces sp.	Actinobacteridae	1.00
luster	: 125 stability	measure : 1.000000		
CHR	NC_011661	D. turgidum	Dictyoglomia	1.00
CHR	NC_011297	D. thermophilum	Dictyoglomia	1.00
luster	: 126 stability	measure : 0.873016		
L I	NC_016975	L. casei	Bacilli	1.00
	NC_004947	L. fermentum	Bacilli	1.00
	NC_002799		Bacilli Da ailli	1.00
- L эт	NC 012542	L. paracasei	Bacilli	1.00
DI.	NC 013544	L. paracasei	Bacilli	1.00
эī.	NC 013789	L. plantarum	Bacilli	0.64
э́Г	NC 011223	L. rhamnosus	Bacilli	1.00
- 1				
luster	: 127 stability	measure : 0.614281		
PL	NC_015561	A. subflavus	Actinobacteridae	0.73
PL	NC_009339	M. gilvum	Actinobacteridae	1.00
PL	NC_008147	Mycobacterium sp.	Actinobacteridae	0.94
	NG OCOTO I	Mucobacterium sp	Actinobacteridae	1.00
PL	NC_008704	<i>mycooucienium</i> sp.		
PL PL	NC_008704 NC_008269	R. jostii	Actinobacteridae	0.71
PL PL	NC_008704 NC_008269	R. jostii	Actinobacteridae	0.71
PL PL	NC-008704 NC-008269 : 128 stability	R. jostii measure : 1.000000	Actinobacteridae	0.71
PL PL eluster PL	NC-008704 NC-008269 : 128 stability NC-006257	R. jostii measure : 1.000000 L. casei	Actinobacteridae Bacilli	0.71
PL PL PL PL PL	NC-008704 NC-008269 : 128 stability NC-006257 NC-006256 NC-004266	R. jostii measure : 1.000000 L. casei L. fermentum	Actinobacteridae Bacilli Bacilli	0.71 1.00 1.00
PL PL PL PL PL PL PL	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_01602	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. casti	Actinobacteridae Bacilli Bacilli Bacilli Bacilli	0.71 1.00 1.00 1.00
PL PL PL PL PL PL PL PL	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_015700	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri L. reuteri	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli	0.71 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012642	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri L. reuteri L. reuteri	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.71 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012642 NC_015876	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri L. reuteri S. pseudomeumoniae	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_015876	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri L. reuteri S. parasanguinis S. pseudopneumoniae	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.71 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL	NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-015700 NC-012642 NC-015876 : 129 stability	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 0.71 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \end{array}$
PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_015876 : 129 stability NC_018878	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	0.71 1.00 1.00 1.00 1.00 1.00 1.00 0.87
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012642 NC-015876 : 129 stability NC-018878 NC-012654	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.87 0.47
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012222 NC-010603 NC-012642 NC-015876 : 129 stability NC-018878 NC-012654 NC-010379	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.87 0.47 1.00
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012222 NC-010603 NC-012642 NC-012642 NC-015876 : 129 stability NC-018878 NC-012654 NC-010379 NC-010418	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia	$\begin{array}{c} 0.71 \\ \hline 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ \hline 0.87 \\ 0.47 \\ 1.00 \\ 0.50 \end{array}$
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012642 NC-018878 : 129 stability NC-018878 NC-012654 NC-010379 NC-010418 NC-010680	R. jostii measure : 1.000000 L. casei L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.71 \\ \hline 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ \hline 0.87 \\ 0.47 \\ 1.00 \\ 0.50 \\ 0.33 \end{array}$
>L	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_012676 : 129 stability NC_018878 NC_012654 NC_010379 NC_010418 NC_010680 NC_017177	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. difficile	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.71 \\ \hline 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ \hline 0.87 \\ 0.47 \\ 1.00 \\ 0.50 \\ 0.33 \\ 0.82 \\ \end{array}$
>L	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_015876 : 129 stability NC_018878 NC_012654 NC_012654 NC_010379 NC_012654 NC_010418 NC_010680 NC_017177	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.71\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.87\\ 0.47\\ 1.00\\ 0.50\\ 0.33\\ 0.82 \end{array}$
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012222 NC-010603 NC-012642 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-010379 NC-010418 NC-010680 NC-01777 : 130 stability	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.71 \\ \hline 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ \hline 0.87 \\ 0.47 \\ 1.00 \\ 0.50 \\ 0.33 \\ 0.82 \\ \hline 0.82 \\ $
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012222 NC-010603 NC-012642 NC-012642 NC-015876 : 129 stability NC-018878 NC-010379 NC-010418 NC-010418 NC-010418 NC-0104717 : 130 stability	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 0.71\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 0.87\\ 0.47\\ 1.00\\ 0.50\\ 0.33\\ 0.82\\ \hline 1.00\\ \end{array}$
>L	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_012642 NC_012654 NC_010878 NC_010680 NC_010418 NC_010680 NC_017177 : 130 stability NC_009472 NC_009472	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. cryptum A. cryptum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.87 0.47 1.00 0.50 0.33 0.82 1.00 nan 1.00
>L >L <th>NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-010680 NC-012654 NC-010680 NC-017177 : 130 stability NC-009472 NC-009474 NC-008474 NC-008474 NC-008474</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum</th> <th>Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Questridia aproteobacteria aproteobacteria aproteobacteria</th> <th>0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.87 0.47 1.00 0.33 0.82 1.00 nan 1.00 1.00</th>	NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-010680 NC-012654 NC-010680 NC-017177 : 130 stability NC-009472 NC-009474 NC-008474 NC-008474 NC-008474	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Questridia aproteobacteria aproteobacteria aproteobacteria	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.87 0.47 1.00 0.33 0.82 1.00 nan 1.00 1.00
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012622 NC-012642 NC-012642 NC-012676 : 129 stability NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-017177 : 130 stability NC-009472 NC-009474 NC-009474 NC-009474	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum A. cryptum A. multivorum A. multivorum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.87 0.47 1.00 0.33 0.82 1.00 nan 1.000 1.00 1.00 1.00 1.0000 1.0000 1.000 1.0000 1.0000 1.00
>L	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012222 NC_010603 NC_012642 NC_012642 NC_015876 : 129 stability NC_018878 NC_012654 NC_010379 NC_010418 NC_012654 NC_010418 NC_010418 NC_010418 NC_009472 NC_009474 NC_009474 NC_009667	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.53 0.82 1.00 nan 1.000
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012222 NC-010603 NC-012642 NC-012642 NC-012642 NC-010379 NC-010418 NC-010418 NC-010418 NC-010418 NC-009472 NC-009474 NC-008691 NC-015187 NC-015187 NC-008676 NC-014009	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum A. cryptum A. cryptum A. multivorum A. multivorum	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.50 0.33 0.82 1.00 nan 1.00
>L	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_012654 NC_012654 NC_010478 NC_010478 NC_010680 NC_017177 : 130 stability NC_009472 NC_009472 NC_009472 NC_009473 NC_015187 NC_015189 NC_015189 NC_015189 NC_006826	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum A. multivorum A. multivorum A. multivorum S. japonicum S. zenophagum	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.50 0.33 0.82 1.00 nan 1.00
>L	NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-015876 : 129 stability NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-010418 NC-010680 NC-017177 : 130 stability NC-009472 NC-009472 NC-009474 NC-009474 NC-009474 NC-005187 NC-015189 NC-006676 NC-006826 NC-008246	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum A. multivorum S. yaponicum S. yaponikuyae	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.33 0.82 1.00 nan 1.00
>L	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012222 NC_010603 NC_012642 NC_012642 NC_012642 NC_015876 : 129 stability NC_018878 NC_012654 NC_010680 NC_012654 NC_010418 NC_010680 NC_017177 : 130 stability NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009476 NC_009474 NC_009474 NC_009476 NC_006676 NC_014009 NC_006826 NC_007505	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum A. multivorum S. japonicum S. yanoikuyae X. campestris	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Questridia Clostridia Clostridia Clostridia aproteobacteria	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.33 0.82 1.00 nan 1.00
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012622 NC-010603 NC-012642 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-009472 NC-009472 NC-009474 NC-009474 NC-008691 NC-015187 NC-015187 NC-015187 NC-006676 NC-014009 NC-006676 NC-006826 NC-008246 NC-007505	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. japonicum S. zenophagum S. yanoikuyae X. campestris	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Qroteobacteria αproteobacteria αproteobacteria </th <th>0.71 1.00</th>	0.71 1.00
>L >L <th>NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_004566 NC_012222 NC_015700 NC_012642 NC_012642 NC_012654 NC_012654 NC_010687 NC_010418 NC_010680 NC_017177 : 130 stability NC_009472 NC_009472 NC_009472 NC_009472 NC_009472 NC_009472 NC_009472 NC_009473 NC_015187 NC_015189 NC_015189 NC_015189 NC_015189 NC_006826 NC_006826 NC_007505 : 131 stability</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. senophagum S. yanoikuyae X. campestris measure : 1.000000</th> <th>Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostr</th> <th>0.71 1.00</th>	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_004566 NC_012222 NC_015700 NC_012642 NC_012642 NC_012654 NC_012654 NC_010687 NC_010418 NC_010680 NC_017177 : 130 stability NC_009472 NC_009472 NC_009472 NC_009472 NC_009472 NC_009472 NC_009472 NC_009473 NC_015187 NC_015189 NC_015189 NC_015189 NC_015189 NC_006826 NC_006826 NC_007505 : 131 stability	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. senophagum S. yanoikuyae X. campestris measure : 1.000000	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostr	0.71 1.00
>L	NC-008704 NC-008269 : 128 stability NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-010418 NC-010418 NC-010418 NC-010680 NC-017177 : 130 stability NC-009472 NC-009472 NC-009474 NC-009474 NC-009474 NC-009474 NC-005187 NC-015189 NC-006826 NC-006826 NC-008246 NC-007505 : 131 stability NC-007104	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. yanoikuyae S. yanoikuyae X. campestris measure : 1.000000 B. cereus	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Qroteobacteria αproteobacteria βaproteobacteria βapr	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.50 0.33 0.82 1.00 nan 1.00
>L	NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-010878 NC-012654 NC-010878 NC-010680 NC-017177 : 130 stability NC-009472 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009475 NC-015189 NC-006826 NC-008246 NC-007505 : 131 stability NC-007104 NC-007104	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum G. oxydans S. japonicum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria βapoteobacteria βapoteobacteria βapoteobacteria βapoteobacteria βapoteobacteria βapoteobacteria βacilli Bacilli Bacilli	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.33 0.82 1.00 nan 1.00
>L >L <th>NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012222 NC_010603 NC_012642 NC_012642 NC_012642 NC_018878 NC_012654 NC_010878 NC_012654 NC_010680 NC_012654 NC_010418 NC_010680 NC_017177 : 130 stability NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009476 NC_009476 NC_006676 NC_014009 NC_006676 NC_006266 NC_007505 : 131 stability NC_007104</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. japonicum S. senophagum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. mycoides</th> <th>Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria β β Βacilli Bacilli Bacilli Bacilli</th> <th>0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.33 0.82 1.00 nan 1.00</th>	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012222 NC_010603 NC_012642 NC_012642 NC_012642 NC_018878 NC_012654 NC_010878 NC_012654 NC_010680 NC_012654 NC_010418 NC_010680 NC_017177 : 130 stability NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009476 NC_009476 NC_006676 NC_014009 NC_006676 NC_006266 NC_007505 : 131 stability NC_007104	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. japonicum S. senophagum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. mycoides	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria β β Βacilli Bacilli Bacilli Bacilli	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.33 0.82 1.00 nan 1.00
>L >L <th>NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012222 NC_010603 NC_012642 NC_012642 NC_010876 : 129 stability NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012657 NC_009472 NC_009472 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009476 NC_006676 NC_014009 NC_006676 NC_006826 NC_006826 NC_007505 : 131 stability NC_005704 NC_005704 NC_004059</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum A. multivorum S. zenophagum S. yanoikuyae X. campestris measure : 1.000000 B. ccreus B. ccreus B. mycoides B. thuringiensis</th> <th>Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Qroteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria β Bacilli Bacilli Bacilli Bacilli Bacilli</th> <th>0.71 1.00</th>	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012222 NC_010603 NC_012642 NC_012642 NC_010876 : 129 stability NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012657 NC_009472 NC_009472 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009476 NC_006676 NC_014009 NC_006676 NC_006826 NC_006826 NC_007505 : 131 stability NC_005704 NC_005704 NC_004059	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum A. multivorum S. zenophagum S. yanoikuyae X. campestris measure : 1.000000 B. ccreus B. ccreus B. mycoides B. thuringiensis	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Qroteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria β Bacilli Bacilli Bacilli Bacilli Bacilli	0.71 1.00
>L	NC_008704 NC_008269 : 128 stability NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_012654 NC_012654 NC_012654 NC_012654 NC_010418 NC_010680 NC_017177 : 130 stability NC_009472 NC_009472 NC_009474 NC_009472 NC_009472 NC_009474 NC_009472 NC_009475 NC_015189 NC_015189 NC_015189 NC_006826 NC_008246 NC_007505 : 131 stability NC_007104 NC_004059 NC_004059 NC_006821	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. zenophagum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. mycoides B. thuringiensis B. thuringiensis	Actinobacteridae Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Clostridia Clostr	0.71 1.00
>L >L <th>NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-010680 NC-017177 : 130 stability NC-009472 NC-009472 NC-009474 NC-009474 NC-009472 NC-009474 NC-009474 NC-009472 NC-009474 NC-005187 NC-0015189 NC-006826 NC-008246 NC-007505 : 131 stability NC-007104 NC-004059 NC-006821 NC-006821 NC-006821 NC-009016</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. yaponicum S. yaponicum S. yaponicum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. cereus B. thuringiensis F. limi</th> <th>Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia aproteobacteria Bacilli Bacill</th> <th>0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.50 0.33 0.82 1.00 nan 1.00 1</th>	NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-010680 NC-017177 : 130 stability NC-009472 NC-009472 NC-009474 NC-009474 NC-009472 NC-009474 NC-009474 NC-009472 NC-009474 NC-005187 NC-0015189 NC-006826 NC-008246 NC-007505 : 131 stability NC-007104 NC-004059 NC-006821 NC-006821 NC-006821 NC-009016	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. yaponicum S. yaponicum S. yaponicum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. cereus B. thuringiensis F. limi	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia aproteobacteria Bacilli Bacill	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.50 0.33 0.82 1.00 nan 1.00 1
>L >L <th>NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012222 NC-010603 NC-012642 NC-015876 : 129 stability NC-018878 NC-012654 NC-012654 NC-012654 NC-012654 NC-010418 NC-010680 NC-017177 : 130 stability NC-009472 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009475 NC-0015189 NC-006826 NC-008246 NC-007104 NC-007104 NC-007104 NC-00704 NC-00704 NC-006821 NC-0019016 NC-01983</th> <th>measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. cryptum A. multivorum G. oxydans S. japonicum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis F. limi</th> <th>Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria βacilli Bacilli Bacilli <tr td=""></tr></th> <th>0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.33 0.82 1.00 nan 1.00</th>	NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012222 NC-010603 NC-012642 NC-015876 : 129 stability NC-018878 NC-012654 NC-012654 NC-012654 NC-012654 NC-010418 NC-010680 NC-017177 : 130 stability NC-009472 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009474 NC-009475 NC-0015189 NC-006826 NC-008246 NC-007104 NC-007104 NC-007104 NC-00704 NC-00704 NC-006821 NC-0019016 NC-01983	measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. cryptum A. multivorum G. oxydans S. japonicum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis F. limi	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Quoteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria αproteobacteria βacilli Bacilli Bacilli <tr td=""></tr>	0.71 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.33 0.82 1.00 nan 1.00
>L >L <th>NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012622 NC_012642 NC_012642 NC_012642 NC_012674 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_009474 NC_009472 NC_009472 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009476 NC_009474 NC_009476 NC_009476 NC_009474 NC_006676 NC_014009 NC_006676 NC_006826 NC_006705 : 131 stability NC_007104 NC_006704 NC_006704 NC_006821 NC_00983 : 132 stability</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum S. sotulinum A. cryptum A. multivorum A. multivorum A. multivorum S. japonicum S. saenophagum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. mycoides B. thuringiensis B. thuringiensis B. thuringiensis F. limi aproteobacteriaacterium</th> <th>Actinobacteridae Bacilli Clostridia Aproteobacteria αproteobacteria αproteobacteria γproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Cytophagia <th>0.71 1.00</th></th>	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012622 NC_012642 NC_012642 NC_012642 NC_012674 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_009474 NC_009472 NC_009472 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009474 NC_009476 NC_009474 NC_009476 NC_009476 NC_009474 NC_006676 NC_014009 NC_006676 NC_006826 NC_006705 : 131 stability NC_007104 NC_006704 NC_006704 NC_006821 NC_00983 : 132 stability	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum S. sotulinum A. cryptum A. multivorum A. multivorum A. multivorum S. japonicum S. saenophagum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. cereus B. mycoides B. thuringiensis B. thuringiensis B. thuringiensis F. limi aproteobacteriaacterium	Actinobacteridae Bacilli Clostridia Aproteobacteria αproteobacteria αproteobacteria γproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Cytophagia <th>0.71 1.00</th>	0.71 1.00
>L >L <th>NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_010680 NC_017177 : 130 stability NC_009472 NC_009472 NC_009472 NC_009474 NC_009472 NC_009472 NC_009474 NC_009472 NC_009474 NC_009475 NC_015187 NC_015189 NC_015189 NC_015189 NC_006826 NC_006826 NC_006826 NC_007505 : 131 stability NC_007104 NC_004059 NC_006821 NC_010983 : 132 stability NC_012020</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. senophagum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis F. limi aproteobacteriaacterium measure : 0.492063</th> <th>Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia aproteobacteria aproteobacteria <</th> <th>0.71 1.00</th>	NC_008704 NC_008269 : 128 stability NC_006257 NC_004566 NC_012222 NC_010603 NC_012642 NC_012642 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_012654 NC_010680 NC_017177 : 130 stability NC_009472 NC_009472 NC_009472 NC_009474 NC_009472 NC_009472 NC_009474 NC_009472 NC_009474 NC_009475 NC_015187 NC_015189 NC_015189 NC_015189 NC_006826 NC_006826 NC_006826 NC_007505 : 131 stability NC_007104 NC_004059 NC_006821 NC_010983 : 132 stability NC_012020	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. botulinum C. difficile measure : 0.857364 A. cryptum A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum S. senophagum S. yanoikuyae X. campestris measure : 1.000000 B. cereus B. thuringiensis B. thuringiensis B. thuringiensis B. thuringiensis F. limi aproteobacteriaacterium measure : 0.492063	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia aproteobacteria aproteobacteria <	0.71 1.00
>L >L <th>NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-010418 NC-010680 NC-017177 : 130 stability NC-009472 NC-009472 NC-009474 NC-009474 NC-009474 NC-009474 NC-005187 NC-015189 NC-006826 NC-006826 NC-006826 NC-007505 : 131 stability NC-004059 NC-006821 NC-01983 : 132 stability NC-01983</th> <th>R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum G. oxydans S. japonicum S. senophagum S. senophagum S. senophagum S. ganoikuyae X. campestris measure : 1.000000 B. cereus B. thuringiensis B. thuringiensis B. thuringiensis F. limi aproteobacteriaacterium measure : 0.492063 L. garvieae L. lactic</th> <th>Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Quoteobacteria aproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli B</th> <th>0.71 1.00</th>	NC-008704 NC-008269 : 128 stability NC-006257 NC-004566 NC-012222 NC-010603 NC-012642 NC-012642 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-012654 NC-010418 NC-010680 NC-017177 : 130 stability NC-009472 NC-009472 NC-009474 NC-009474 NC-009474 NC-009474 NC-005187 NC-015189 NC-006826 NC-006826 NC-006826 NC-007505 : 131 stability NC-004059 NC-006821 NC-01983 : 132 stability NC-01983	R. jostii measure : 1.000000 L. casei L. fermentum L. paracasei L. reuteri S. parasanguinis S. pseudopneumoniae measure : 0.534934 B. thuringiensis C. botulinum C. difficile measure : 0.857364 A. cryptum A. multivorum A. multivorum A. multivorum A. multivorum G. oxydans S. japonicum S. senophagum S. senophagum S. senophagum S. ganoikuyae X. campestris measure : 1.000000 B. cereus B. thuringiensis B. thuringiensis B. thuringiensis F. limi aproteobacteriaacterium measure : 0.492063 L. garvieae L. lactic	Actinobacteridae Bacilli Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia Quoteobacteria aproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli B	0.71 1.00

Annexe D.Résultats du clustering de V^R par INFOMAP

\mathbf{PL}	NC_013657	L. lactis	Bacilli	1.00	\mathbf{PL}	NC_017204	B. thuringiensis	Bacilli	1.00
PL	NC_015902	L. lactis	Bacilli	nan	\mathbf{PL}	NC_018694	B. thuringiensis	Bacilli	1.00
PL	NC_015864	L. lactis	Bacilli	1.00	\mathbf{PL}	NC_018689	B. thuringiensis	Bacilli	0.00
PL	NC_017496	L. lactis	Bacilli	1.00	\mathbf{PL}	NC_018487	B. thuringiensis	Bacilli	1.00
	1					I			
cluster :	133 stability	measure : 1.000000			cluster	: 145 stability	measure : 0.675833		
CHR	NC_010503	U. parvum	Mollicutes	1.00	PL	NC_003042	C. perfringens	Clostridia	0.47
CHR	NC_002162	U. parvum	Mollicutes	1.00	\mathbf{PL}	NC_006872	C. perfringens	Clostridia	1.00
CHR	NC-011374	U. urealuticum	Mollicutes	1.00	PL	NC-008263	C. perfringens	Clostridia	0.43
		0.1			PL	NC 008264	C. perfringens	Clostridia	1.00
cluster :	134 stability	measure : 0.607222			PL	NC 015427	C hotulinum	Clostridia	1.00
PL.	NC 011061	P aestuarii	Chlorobia	0.33		1101010121	01 000000000	crossinaia	1.00
DI	NC 005702	T thermorphilus	Deipegegi	1.00	clustor	• 146 stability	$m_{02} = 0.567321$		
	NC 006462	T thermophilus	Deinococci	1.00	DI	NC 017479	L and	Da elli	0.69
	NC 007100	T thermophilus	Deinessei	1.00		NC 001670	L. unigiovorus	Dacilli	1.00
PL	NC_007190	1. thermophilus	Deinococci	1.00	PL	NC_001670	L. aelorueckii	Bacilli	1.00
PL	NC_015716	Z. mobilis	α proteobacteria	0.25	PL	NC_015598	L. kefiranofaciens	Bacilli	1.00
PL	NC_018148	Z. mobilis	α proteobacteria	1.00	PL	NC-005322	S. thermophilus	Bacilli	1.00
PL	NZ_CM00136	59 Desulfovibrio sp.	∂proteobacteria	0.62	PL	NC_010859	S. thermophilus	Bacilli	1.00
PL	NC_007515	G. metallireducens	∂proteobacteria	0.12	PL	NC_013281	bacterium	Unidentified	1.00
PL	NC_009794	C. koseri	γ proteobacteria	0.56	- · ·				
PL	NC_016747	Oceanimonas sp.	γ proteobacteria	0.56	cluster	: 147 stability	measure : 1.000000		
					\mathbf{PL}	NC_012674	P. fluorescens	γ proteobacteria	1.00
cluster :	135 stability	measure : 0.532552			\mathbf{PL}	NC_003350	P. putida	γ proteobacteria	1.00
PL	NC_004535	C. glutamicum	Actinobacteridae	1.00	\mathbf{PL}	NC_004999	P. putida	γ proteobacteria	1.00
CHR	NC_013260	Cand. Methylomirabilis	Unidentified	nan	\mathbf{PL}	NC_007926	P. putida	γ proteobacteria	1.00
					\mathbf{PL}	NC_016644	Pseudomonas sp.	γ proteobacteria	1.00
cluster :	136 stability	measure : 0.680895							
PL	NC_009471	A. cryptum	α proteobacteria	0.79	cluster	: 148 stability	measure : 0.770238		
PL	NC_013356	Z. mobilis	α proteobacteria	0.25	\mathbf{PL}	NC_010539	Cyanothece sp.	Chroobacteria	1.00
\mathbf{PL}	NC_013785	Z. mobilis	α proteobacteria	1.00	\mathbf{PL}	NC_011730	Cyanothece sp.	Chroobacteria	1.00
\mathbf{PL}	NC_017184	Z. mobilis	α proteobacteria	0.50	\mathbf{PL}	NC_011733	Cyanothece sp.	Chroobacteria	1.00
PL	NC_011667	Thauera sp.	β proteobacteria	0.91	\mathbf{PL}	NC_011734	Cyanothece sp.	Chroobacteria	1.00
PL	NC_004632	P. syringae	γ proteobacteria	0.14	\mathbf{PL}	NC_014503	Cyanothece sp.	Chroobacteria	0.77
PL	NC_007714	S. glossinidius	γproteobacteria	1.00	\mathbf{PL}	NC_014535	Cyanothece sp.	Chroobacteria	1.00
PL	NC_007184	S. glossinidius	γ proteobacteria	0.90	\mathbf{PL}	NC_010478	Synechococcus sp.	Chroobacteria	1.00
PL	NC_007185	S. glossinidius	γproteobacteria	1.00	\mathbf{PL}	NC_011760	M. chloromethanicum	α proteobacteria	0.64
		0	1.		\mathbf{PL}	NC_008758	P. naphthalenivorans	β proteobacteria	0.23
cluster :	137 stability	measure : 0.805714				I	•	, <u>,</u>	1
PL	NC_009999	S. baltica	γproteobacteria	1.00	cluster	: 149 stability	measure : 1.000000		
\mathbf{PL}	NC_011665	S. baltica	γproteobacteria	1.00	PL	NC_017213	B. thuringiensis	Bacilli	1.00
PL	NC_011422	S. enterica	vproteobacteria	1.00	PL	NC_018885	B. thuringiensis	Bacilli	1.00
\mathbf{PL}	NC_005128	V. vulnificus	γproteobacteria	1.00	\mathbf{PL}	NC_007202	B. thuringiensis	Bacilli	1.00
PL	NC-009701	V. vulnificus	vproteobacteria	1.00	PL	NC_017197	B. thuringiensis	Bacilli	1.00
PL	NC 009702	V. vulnificus	vproteobacteria	1.00	PL	NC 002108	B. thurinaiensis	Bacilli	1.00
PL	NC 009703	V vulnificus	vproteobacteria	1.00					
PL	NC 010112	Vihria sp	vproteobacteria	1.00	cluster	• 150 stability	measure : 0 559444		
	11010111	100100 001	procobaccona	1 100	orabuor				
					PL	NC 009353	M arymhiswaldense	oprotoobactoria	1.00
cluster .	138 stability	measure $\cdot 0.799315$		I	PL	NC_009353	M. gryphiswaldense M. marinus	α proteobacteria	1.00
cluster :	138 stability	measure : 0.799315	aprotosbactoria		PL CHR	NC_009353 NC_008576	M. gryphiswaldense M. marinus	α proteobacteria Unidentified	$ \begin{array}{c} 1.00 \\ 0.67 \end{array} $
cluster : PL PI	138 stability NC_010721	measure : 0.799315 M. populi	α proteobacteria	0.92	PL CHR	NC_009353 NC_008576	M. gryphiswaldense M. marinus	α proteobacteria Unidentified	$1.00 \\ 0.67$
cluster : PL PL DI	138 stability NC_010721 NC_009433	measure : 0.799315 M. populi R. sphaeroides	α proteobacteria α proteobacteria	0.92	PL CHR cluster	NC_009353 NC_008576 : 151 stability	M. gryphiswaldense M. marinus measure : 0.629402	α proteobacteria Unidentified	1.00 0.67
cluster : PL PL PL PL PI	138 stability NC_010721 NC_009433 NC_018830 NC_007617	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multifermic	α proteobacteria α proteobacteria β proteobacteria	0.92 1.00 0.87 1.00	PL CHR cluster PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_000066	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae	α proteobacteria Unidentified γ proteobacteria	1.00 0.67
cluster : PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006120	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. pseubarabila	α proteobacteria α proteobacteria β proteobacteria β proteobacteria	0.92 1.00 0.87 1.00 1.00	PL CHR cluster PL PL DI	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffi	αproteobacteria Unidentified γproteobacteria γproteobacteria	1.00 0.67 1.00 0.50
cluster : PL PL PL PL PL PL PI	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_011314	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmowicida	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria δ proteobacteria	0.92 1.00 0.87 1.00 1.00 0.95	PL CHR cluster PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria	1.00 0.67 1.00 0.50 1.00
cluster : PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_011314	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. salarozbij	α proteobacteria α proteobacteria β proteobacteria β proteobacteria δ proteobacteria γ proteobacteria	0.92 1.00 0.87 1.00 1.00 0.95 0.91	PL CHR cluster PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_018220	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thicell bishchi are	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 0.67\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.04 \end{array} $
cluster : PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_011314 NC_009779 NC_012324	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. twicsopic	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.02 \end{array}$	PL CHR cluster PL PL PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp.	α proteobacteria Unidentified γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$ \begin{array}{c} 1.00\\ 0.67\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.94\\ \end{array} $
cluster : PL PL PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_07617 NC_006139 NC_011314 NC_009779 NC_013284	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. osomoinoos	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.04 \end{array}$	PL CHR Cluster PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp.	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$ \begin{array}{c} 1.00\\ 0.67\\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.94\\ \end{array} $
cluster : PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_011314 NC_009779 NC_013284 NC_009739	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa Y. eit	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.94 \\ 0.02 \end{array}$	PL CHR cluster PL PL PL PL PL cluster	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.67 1.00 0.50 1.00 1.00 0.94
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_011314 NC_009779 NC_013284 NC_009739 NC_005240	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri V. rettic	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.94 \\ 0.93 \\ 0.02 \end{array}$	PL CHR cluster PL PL PL PL PL Cluster CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 0.67 1.00 1.00 1.00 0.94
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.011314 NC.009779 NC.013284 NC.009739 NC.005240 NC.005814	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.94 \\ 0.93 \\ 0.93 \\ 0.93 \end{array}$	PL CHR cluster PL PL PL PL PL CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes	1.00 0.67 1.00 1.00 1.00 0.94 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_01314 NC_009779 NC_013284 NC_009739 NC_005240 NC_005814	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.93 \\ 0.93 \\ 0.93 \end{array}$	PL CHR PL PL PL PL Cluster CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes	$\begin{array}{c} 1.00\\ 0.67\\ \hline \\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.94\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_01314 NC_009779 NC_013284 NC_009739 NC_005240 NC_005814 139 stability	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879	α proteobacteria α proteobacteria β proteobacteria β proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.94 \\ 0.93 \\ 0.93 \\ \end{array}$	PL CHR PL PL PL PL PL CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes	$\begin{array}{c} 1.00\\ 0.67\\ \hline \\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 0.94\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL Cluster : PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_011314 NC_009779 NC_013284 NC_009739 NC_005240 NC_005814 139 stability NC_015171 NC_01400	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. hyicus	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.94 \\ 0.93 \\ 0.93 \\ \hline \end{array}$	PL CHR PL PL PL PL PL PL CHR CHR CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes	1.00 0.67 1.00 1.00 1.00 1.00 1.00 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.007617 NC.006139 NC.011314 NC.009739 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.004986	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. rumingatium	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.93 \\ 0.93 \\ 0.93 \\ \hline \end{array}$	PL CHR CLUSTER PL PL PL PL PL CHR CHR CHR CHR CHR CLUSTER PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis measure : 0.476937 E. faccium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli	$ \begin{array}{c} 1.00\\ 0.67\\ \hline 1.00\\ 0.50\\ 1.00\\ 0.94\\ \hline 1.00\\ 1.00\\ \hline $
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.01314 NC.009779 NC.013284 NC.009739 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.015171 NC.019149 NC.004986	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. muninantium	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria Bacilli Bacilli Negativicutes	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.93 \\ 0.93 \\ 0.93 \\ \hline \end{array}$	PL CHR CLUSTER PL PL PL CLUSTER CHR CHR CHR CHR CHR CHR CHR CHR CHR CH	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_016829 NC_017519 : 153 stability NC_010291 NC_010330	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis measure : 0.476937 E. faecium E. faecium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli Bacilli	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_01314 NC_009779 NC_013284 NC_009739 NC_005240 NC_005240 NC_005814 139 stability NC_015171 NC_019149 NC_004986 NC_013776	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium	lpha proteobacteria lpha proteobacteria eta proteobacteria δ proteobacteria γ proteobacteria	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ \hline \end{array}$	PL CHR PL PL PL PL PL CHR CHR CHR CHR CHR CHR CHR CHR PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability NC_010291 NC_010330 NC_017024	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faecium E. faecium E. faecium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 0.67\\ \hline \\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC_010721 NC_009433 NC_018830 NC_007617 NC_006139 NC_011314 NC_009779 NC_013284 NC_009739 NC_005814 139 stability NC_015171 NC_019149 NC_004986 NC_013776	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium	α proteobacteria α proteobacteria β proteobacteria β proteobacteria γ proteobacteria	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.01\\$	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_016829 NC_017519 : 153 stability NC_010291 NC_010330 NC_017024	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faecium E. faecium E. faecium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 0.67\\ \hline \\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.01314 NC.006139 NC.011314 NC.009739 NC.00739 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium measure : 0.077862 M. piecephila	α proteobacteria α proteobacteria β proteobacteria β proteobacteria γ proteobac	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.93 \\ 0.93 \\ 0.93 \\ \hline 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ 1.00 \\ \hline \end{array}$	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_016829 NC_016829 NC_017519 : 153 stability NC_010291 NC_010291 NC_017024 : 154 stability	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faecium E. faecium E. faecium E. faecium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 0.67\\ \hline \\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ \hline \\ 1.00\\ 1.00\\ \hline \\ 1.00\\ 1.00\\ \hline \\ 1.00\\ \hline \\ 1.00\\ \hline \\ 1.00\\ \hline \\ 1.00\\ \hline \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.006139 NC.01314 NC.009779 NC.013284 NC.005240 NC.005240 NC.015171 NC.015171 NC.013776 : 140 stability NC.016751 NC.016751	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium measure : 0.077862 M. piezophila P. mobilis	α proteobacteria α proteobacteria β proteobacteria δ proteobacteria γ proteobacteria Thermotogae Thermotogae	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ \hline \end{array}$	PL CHR CLUSTER PL PL PL PL CLUSTER CHR CHR CHR CHR CHR CHR CHR CHR CHR CH	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_016829 NC_017519 : 153 stability NC_010291 NC_010300 NC_017024 : 154 stability NC_007930 NC_07930	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.01830 NC.006139 NC.01314 NC.009779 NC.013284 NC.005240 NC.015171 NC.015171 NC.019149 NC.019149 NC.019176 140 stability NC.016751 NC.010003	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium measure : 0.077862 M. piezophila P. mobilis	$\begin{array}{c} \alpha \mathrm{proteobacteria} \\ \alpha \mathrm{proteobacteria} \\ \beta \mathrm{proteobacteria} \\ \beta \mathrm{proteobacteria} \\ \delta \mathrm{proteobacteria} \\ \gamma \mathrm{proteobacteria} \\ \lambda \mathrm{proteobacteria} \\ proteobacte$	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\ 0.93\\ 0.93\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_016829 NC_016829 NC_016829 NC_017519 : 153 stability NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_017499 NC_017676	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. salivarius L. salivarius L. salivarius	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli Bacilli Bacilli	1.00 0.67 1.00 0.50 1.00 0.94 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.011314 NC.009739 NC.013284 NC.009739 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability NC.016751 NC.010003	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis	$\begin{array}{c} \alpha proteobacteria\\ \alpha proteobacteria\\ \beta proteobacteria\\ \beta proteobacteria\\ \gamma proteobacteria\\ li \\ Bacilli\\ Bacilli\\ Negativicutes\\ Negativicutes\\ Thermotogae\\ Thermotogae\\ \end{array}$	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.00\\$	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_007930 NC_017499 NC_017076	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faecium E. faecium E. faecium E. faecium E. faecium E. salivarius L. salivarius S. ruminantium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.006139 NC.01314 NC.00379 NC.005240 NC.005240 NC.015171 NC.015171 NC.013776 140 stability NC.01003 141 stability	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393	α proteobacteria α proteobacteria β proteobacteria β proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria γ proteobacteria Thermotogae Thermotogae	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.$	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_016829 NC_017519 : 153 stability NC_010291 NC_010291 NC_01030 NC_017024 : 154 stability NC_007930 NC_017076 NC_017078	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria % Mollicutes Mollicutes Mollicutes Bacilli Bacinteria Bacilli	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018300 NC.007617 NC.006139 NC.013144 NC.009779 NC.013284 NC.009739 NC.005240 NC.005240 NC.005240 NC.015171 NC.015171 NC.019149 NC.013776 140 stability NC.016751 NC.016751 NC.019003 141 stability NC.009488 NC.009488 NC.009488 NC.009488 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.0$	PL CHR cluster PL PL PL Cluster CHR CHR CHR CHR CHR CHR CHR CHR CHR PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010330 NC.017024 : 154 stability NC.007930 NC.017076 NC.017078	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faecium E. faecium E. faecium E. faecium E. faecium S. ruminantium	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Mollicutes Mollicutes Mollicutes Bacilli Bacilli <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.007617 NC.006139 NC.01314 NC.009779 NC.013284 NC.005240 NC.005240 NC.015171 NC.016751 NC.010003 141 stability NC.009488 NC.010793	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi	$\begin{array}{c} \alpha \mathrm{proteobacteria} \\ \alpha \mathrm{proteobacteria} \\ \beta \mathrm{proteobacteria} \\ \beta \mathrm{proteobacteria} \\ \delta \mathrm{proteobacteria} \\ \gamma \mathrm{proteobacteria} \\ \end{array}$	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\ 0.93\\ 0.93\\ 0.93\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\$	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_016829 NC_016829 NC_016829 NC_016829 NC_016829 NC_016829 NC_016829 NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_017076 NC_017076 NC_017078 : 155 stability	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium M. asure : 0.476937	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.00779 NC.01314 NC.009779 NC.013284 NC.005240 NC.005240 NC.015171 NC.013149 NC.005240 NC.005814 139 stability NC.015171 NC.004986 NC.016751 NC.016751 NC.00003 141 stability NC.009488 NC.010793 142 stability	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium Measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi M. piezophila	$\begin{array}{c} \alpha proteobacteria\\ \alpha proteobacteria\\ \beta proteobacteria\\ \beta proteobacteria\\ \gamma proteobacteria\\ \end{array}$	$\begin{array}{c} 0.92\\ 1.00\\ 0.87\\ 1.00\\ 1.00\\ 0.95\\ 0.91\\ 0.93\\$	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR PL PL PL PL PL PL PL Cluster	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_017076 NC_017078 : 155 stability NC_015153	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium M. suis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.007617 NC.006139 NC.011314 NC.009779 NC.012844 NC.005240 NC.005240 NC.005240 NC.005240 NC.015171 NC.0151711 NC.0151711 NC.0151711 NC.013776 140 stability NC.010003 141 stability NC.009488 NC.010793 142 stability 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi	$\begin{array}{c} \alpha proteobacteria \\ \alpha proteobacteria \\ \beta proteobacteria \\ \beta proteobacteria \\ \delta proteobacteria \\ \gamma proteobacteria \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL CHR CHR CHR	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_016829 NC_017519 : 153 stability NC_010291 NC_010291 NC_010291 NC_017024 : 154 stability NC_007930 NC_017076 NC_017078 : 155 stability NC_015153 NC_015155 NC_015155	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium M. sulvarius S. ruminantium S. ruminantium measure : 0.476937 M. suis M. suis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.01314 NC.009779 NC.013284 NC.009739 NC.005240 NC.005240 NC.005240 NC.005240 NC.005241 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.009488 NC.010793 142 stability NC.01843 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium Measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. bifleza	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	$\begin{array}{c} 0.92 \\ 1.00 \\ 0.87 \\ 1.00 \\ 0.95 \\ 0.91 \\ 0.93 \\ 0.$	PL CHR CHR PL PL PL PL Cluster CHR CHR CHR CHR PL PL PL PL PL PL Cluster CHR CHR CHR CHR CHR CHR	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010330 NC.017024 : 154 stability NC.007930 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.015155 NC.018149	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faecium E. faecium E. faecium E. faecium E. faecium E. faecium S. ruminantium measure : 0.476937 M. suis M. suis M. wenyonii	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.01830 NC.006139 NC.01314 NC.009779 NC.013284 NC.009739 NC.005240 NC.005240 NC.005241 139 stability NC.015171 NC.015171 NC.015171 NC.015171 NC.019149 NC.013776 140 stability NC.010003 141 stability NC.009488 NC.010793 142 stability NC.009483 NC.010843 NC.01845 NC.01845 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. bifleza L. bifleza	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability NC_010291 NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_017076 NC_017076 NC_017078 : 155 stability NC_015155 NC_015155 NC_018149	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium E. salivarius S. ruminantium S. ruminantium S. ruminantium M. suis M. suis M. wenyonii	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.013830 NC.007617 NC.006139 NC.01314 NC.009739 NC.005240 NC.005240 NC.005240 NC.005240 NC.005240 NC.005240 NC.015171 NC.015171 NC.019149 NC.00486 NC.013776 140 stability NC.010003 141 stability NC.009488 NC.010793 142 stability NC.010843 NC.010843 NC.010845 NC.008509	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium Measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. biflexa L. borgpetersenii	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 1.00 0.93 0.94 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability NC_010291 NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_017078 : 155 stability NC_015155 NC_015155 NC_018149 : 156 stability	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis measure : 0.476937 E. faecium E. faecium E. faecium E. faecium M. salivarius J. salivarius J. salivarius S. ruminantium S. ruminantium M. suis M. suis M. suis M. suis M. suis (0.734444	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %policutes Mollicutes %policutes %policutes Mollicutes Mollicutes Mollicutes	$\begin{array}{c} 1.00\\ 0.67\\ \hline \\ 1.00\\ 0.50\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ 0.43\\ 0.50\\ \hline \\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.006139 NC.01314 NC.009779 NC.013284 NC.005240 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.013776 140 stability NC.010003 141 stability NC.0109488 NC.010793 142 stability NC.010843 NC.010845 NC.008509 NC.008501 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi Measure : 0.943750 L. biflexa L. borgpetersenii L. borgpetersenii L. borgpetersenii	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.018448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.017024 : 154 stability NC.007930 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.018149 : 156 stability NC.015166	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium M. salivarius S. ruminantium S. ruminantium measure : 0.476937 M. suis M. suis M. suis M. suis M. wenyonii measure : 0.734444 B. salanitronis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.01314 NC.009779 NC.013284 NC.009739 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.015171 NC.019149 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.016751 NC.0109488 NC.010793 142 stability NC.01843 NC.010843 NC.008509 NC.008511 NC.003434 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. biflexa L. biflexa L. borgpetersenii L. interrogans	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010300 NC.017024 : 154 stability NC.007930 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.015155 NC.015166 NC.015166 NC.015168	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium S. ruminantium M. suis M.	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.011314 NC.009779 NC.013284 NC.005240 NC.005240 NC.005240 NC.005240 NC.005240 NC.0105171 NC.015171 NC.013776 140 stability NC.016751 NC.0109488 NC.010793 141 stability NC.009488 NC.010793 142 stability NC.00843 NC.00843 NC.008509 NC.004343 NC.004343 NC.004343 NC.0058211 NC.005824 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. biflexa L. borgpetersenii L. interrogans L. interrogans	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 1.00	PL CHR CLuster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_016829 NC_016829 NC_016829 NC_016829 NC_016829 NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_017076 NC_017076 NC_017076 NC_017078 : 155 stability NC_015153 NC_018149 : 156 stability NC_015168 NC_011564	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium S. ruminantium S. ruminantium M. suis M. suis M. suis M. suis M. suis M. suis M. suinis B. salanitronis B. salanitronis Cand. Azobacteroides	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %mollicutes %mollicutes Mollicutes Mollicutes %mollicutes	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.01314 NC.006139 NC.01314 NC.009739 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability NC.016751 NC.01003 141 stability NC.010793 142 stability NC.010793 142 stability NC.010843 NC.010845 NC.008509 NC.008501 NC.008524 NC.005824 NC.017552	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. bifleza L. interrogans L. interrogans L. interrogans	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.010291 NC.017024 : 154 stability NC.007930 NC.017024 : 155 stability NC.017078 : 155 stability NC.015153 NC.015155 NC.018149 : 156 stability NC.015166 NC.011564 NC.011336	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. salivarius S. ruminantium S. ruminantium M. suis M. suis M. suis M. suis M. suis M. wenyonii measure : 0.734444 B. salanitronis Cand. Azobacteroides C. canimorsus	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.010721 NC.009433 NC.010721 NC.009433 NC.006139 NC.01314 NC.00379 NC.013284 NC.005240 NC.005240 NC.005240 NC.005240 NC.015171 NC.015171 NC.019149 NC.019149 NC.013776 140 stability NC.010003 141 stability NC.0109488 NC.010793 142 stability NC.00845 NC.008509 NC.008509 NC.004343 NC.004343 NC.007552 NC.018021	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piczophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. bifleza L. bifleza L. bifleza L. biflezas L. interrogans L. interrogans L. interrogans L. interrogans T. parva	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR CHR PL PL PL PL Cluster CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.01030 NC.017024 : 154 stability NC.007930 NC.017076 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.018149 : 156 stability NC.01536 NC.011566 NC.011366 NC.011336 NC.011414	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium M. suliarius S. ruminantium measure : 0.476937 M. suis M. suinis M. suis M. suis M	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.01830 NC.007617 NC.006139 NC.01314 NC.009779 NC.013284 NC.009739 NC.005240 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.04986 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.009488 NC.010793 142 stability NC.010843 NC.010843 NC.008509 NC.008511 NC.004324 NC.0018021	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. biflexa L. biflexa L. biflexa L. biflexa L. biflexa L. interrogans L. interrogans L. interrogans T. parva	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.010300 NC.017024 : 154 stability NC.007930 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.015155 NC.015166 NC.01566 NC.011564 NC.011360 NC.011361 NC.011414	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. sufarius L. salivarius L. salivarius L. salivarius S. ruminantium M. suis M. suis M. suis M. suis M. wenyonii measure : 0.73444 B. salanitronis B. salanitronis Cand. Azobacteroides C. canimorsus O. rhinotracheale	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.007617 NC.006139 NC.011314 NC.009779 NC.013284 NC.005240 NC.005240 NC.005240 NC.015171 NC.015171 NC.0149486 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.010793 142 stability NC.00843 NC.008438 NC.010843 NC.008509 NC.004343 NC.004343 NC.004343 NC.004321 143 stability 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. biflexa L. biflexa L. biflexa L. biflexa L. biflexa L. biflexa L. biflexa L. interrogans L. interrogans L. interrogans T. parva measure : 0.315348	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.010291 NC.010330 NC.017024 : 154 stability NC.017076 NC.017076 NC.017076 NC.017078 : 155 stability NC.015155 NC.018149 : 156 stability NC.015166 NC.011564 NC.011564 NC.011344 : 157 stability	M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium M. salivarius J. salivarius J. salivarius S. ruminantium S. ruminantium M. suis M. suoponii measure : 0.734444 B. salanitronis B. salanitronis Cand. Azobacteroides C. canimorsus O. rhinotracheale measure : 0.931944	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli %proteobacteria %proteo	1.00 0.67 1.00 0.50 1.00 0.94 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.01314 NC.006139 NC.01314 NC.009779 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability NC.016751 NC.010043 NC.0109488 NC.010793 141 stability NC.010843 NC.010843 NC.010843 NC.008509 NC.008509 NC.008511 NC.008524 NC.005824 NC.018021 143 stability	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium Measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi C. tsuftexa L. biflexa L. interrogans L. interrogans T. parva measure : 0.315348 N. farcinica	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.018448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.010291 NC.017024 : 154 stability NC.007930 NC.017076 NC.017076 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.018149 : 156 stability NC.015166 NC.011566 NC.011564 NC.011336 NC.011414 : 157 stability NC.008265	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. salivarius S. ruminantium measure : 0.434587 L. salivarius S. ruminantium M. suisantium M. suis M. suorhinis Cand. Azobacteroides C. canimorsus O. rhinotracheale measure : 0.931944 C. perfringens	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>1.00 0.67 1.00 0.50 1.00</th>	1.00 0.67 1.00 0.50 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.01314 NC.009739 NC.005240 NC.005240 NC.005240 NC.015171 NC.015171 NC.015171 NC.019149 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.016751 NC.0100388 NC.010793 142 stability NC.00845 NC.008509 NC.008511 NC.008524 NC.017552 NC.018021 143 stability NC.006363 NZ.CM00091 NZ.CM00091 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piczophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi Measure : 0.943750 L. bifleza L. bifleza L. bifleza L. bifleza L. bifleza L. bifleza L. bifleza L. interrogans L. interrogans T. parva measure : 0.315348 N. farcinica 4.5. clavuligerus	$\begin{array}{c} \alpha \text{proteobacteria} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \beta \text{proteobacteria} \\ \delta \text{proteobacteria} \\ \gamma \text{proteobacteria} \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.01030 NC.017024 : 154 stability NC.017076 NC.017076 NC.017078 : 155 stability NC.015155 NC.015155 NC.015166 NC.011564 NC.011364 NC.011364 NC.011414 : 157 stability NC.008265 NC.007772	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium measure : 0.476937 M. suis S. ruminantium measure : 0.476937 M. suis M. suis	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>1.00 0.67 1.00 0.50 1.00 0.94 1.00</th>	1.00 0.67 1.00 0.50 1.00 0.94 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.01830 NC.007617 NC.006139 NC.01314 NC.009779 NC.013284 NC.005240 NC.005240 NC.005240 NC.015171 NC.015171 NC.019149 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.009488 NC.010793 142 stability NC.008433 NC.008509 NC.008511 NC.00821 143 stability NC.006363 NZ.CM00091 NZ.CM00101 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. biflexa L. biflexa L. biflexa L. biflexa L. biflexa L. interrogans L. interrogans L. interrogans T. parva measure : 0.315348 N. farcinica 4. S. clavuligerus	αproteobacteria βproteobacteria βproteobacteria βproteobacteria γproteobacteria β β β γproteobacteria γproteobacteria γproteobacteria γproteobacteria β γproteobacteria γproteobacteria φ γ γ γ γ γ γ γ γ γ γ γ γ γ γ γ γ γ γ <	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_013300 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability NC_010291 NC_010291 NC_010291 NC_010300 NC_017024 : 154 stability NC_007930 NC_017076 NC_017076 NC_017076 NC_017078 : 155 stability NC_015155 NC_015155 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015168 NC_015166 NC_011336 NC_011414 : 157 stability NC_008265 NC_007772 NC_007773	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis measure : 0.476937 E. faccium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium S. ruminantium M. suis M. suis M	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria Bacilli Bacilli <td< th=""><th>$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$</th></td<>	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$
cluster : PL PL PL PL PL PL PL PL PL PL PL PL Cluster : CHR CHR CHR CHR CHR CHR CHR CHR CHR CHR	 138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.007617 NC.006139 NC.01314 NC.009779 NC.013284 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.015171 NC.013776 140 stability NC.016751 NC.0109488 NC.010793 141 stability NC.010843 NC.008843 NC.0088511 NC.004843 NC.008524 NC.017552 NC.018021 143 stability NC.006363 NZ.CM00091 NZ.CM00101 NZ.006912 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi D. tsutsugamushi D. tsuftexa L. biftexa L. biftexa L. biftexa L. biftexa L. biftexa L. interrogans L. interrogans L. interrogans T. parva measure : 0.315348 N. farcinica 14 S. clavuligerus Streptomyces sp.	αproteobacteria αproteobacteria βproteobacteria βproteobacteria γproteobacteria β β β β β β β β β β β β β β β γ γ β γ γ γ γ γ γ γ γ γ γ γ γ γ <t< th=""><th>0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00</th><th>PL CHR cluster PL PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL</th><th>NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.0145448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.010330 NC.017024 : 154 stability NC.017076 NC.017076 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.018149 : 156 stability NC.015166 NC.011564 NC.011564 NC.011564 NC.011564 NC.011564 NC.011564 NC.011565 NC.007772 NC.007773 NC.010937</th><th>M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium S. ruminantium M. suis M. wenyonii measure : 0.734444 B. salanitronis B. salanitronis Cand. Azobacteroides C. canimorsus O. rhinotracheale measure : 0.931944 C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens</th><th>αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %mollicutes Mollicutes Mollicutes</th><th>1.00 0.67 1.00 0.50 1.00 0.94 1.00</th></t<>	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.0145448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.010330 NC.017024 : 154 stability NC.017076 NC.017076 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.018149 : 156 stability NC.015166 NC.011564 NC.011564 NC.011564 NC.011564 NC.011564 NC.011564 NC.011565 NC.007772 NC.007773 NC.010937	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis M. hyorhinis E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium S. ruminantium S. ruminantium M. suis M. wenyonii measure : 0.734444 B. salanitronis B. salanitronis Cand. Azobacteroides C. canimorsus O. rhinotracheale measure : 0.931944 C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %mollicutes Mollicutes	1.00 0.67 1.00 0.50 1.00 0.94 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.018830 NC.01314 NC.006139 NC.01314 NC.009779 NC.005240 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability NC.010845 NC.010003 141 stability NC.010843 NC.010845 NC.008845 NC.008509 NC.008509 NC.008511 NC.008524 NC.008524 NC.008221 143 stability NC.006363 NZ.CM00010 NZ.CM00101 NC.006912	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piczophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi D. tsuflexa L. biflexa L. biflexa L. biflexa L. biflexa L. biflexa L. interrogans L. interrogans L. interrogans T. parva measure : 0.315348 N. farcinica 4.5. clavuligerus Streptomyces sp.	$\begin{array}{c} \alpha proteobacteria \\ \alpha proteobacteria \\ \beta proteobacteria \\ \beta proteobacteria \\ \delta proteobacteria \\ \gamma proteobacteria \\ \end{array}$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.018448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.01030 NC.017024 : 154 stability NC.007930 NC.017076 NC.017078 : 155 stability NC.015153 NC.015155 NC.018149 : 156 stability NC.01566 NC.011564 NC.011564 NC.011564 NC.011564 NC.011565 NC.007773 NC.00937 NC.001412	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. facium E. faccium E. faccium E. faccium E. faccium E. faccium S. ruminantium Measure : 0.434587 L. salivarius S. ruminantium M. suis M. wenyonii measure : 0.734444 B. salanitronis B. salanitronis C. cand. Azobacteroides C. canimorsus O. rhinotracheale measure : 0.931944 C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria </th <th>1.00 0.67 1.00 0.50 1.00</th>	1.00 0.67 1.00 0.50 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	 138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.013284 NC.009739 NC.005240 NC.005240 NC.005240 NC.005814 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.0109488 NC.010793 142 stability NC.008451 NC.008509 NC.008511 NC.004343 NC.008524 NC.017552 NC.018021 143 stability NC.006363 NZ.CM0091 NZ.CM00101 NC.006912 144 stability 	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi measure : 0.943750 L. biflexa L. biflexa L. biflexa L. biflexa L. biflexa L. interrogans L. interrogans L. interrogans T. parva measure : 0.315348 N. farcinica 14 S. clavuligerus 9 S. clavuligerus Streptomyces sp. measure : 0.893333	$\begin{array}{c} & \alpha \text{proteobacteria} \\ & \alpha \text{proteobacteria} \\ & \beta \text{proteobacteria} \\ & \beta \text{proteobacteria} \\ & \gamma \text{proteobacteria} \\ \\ & \gamma \text{proteobacteria} \\ \\ & \gamma \text{proteobacteria} \\ \\ & \text{Bacilli} \\ & Bacilli$	0.92 1.00 0.87 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.009353 NC.008576 : 151 stability NC.014544 NC.009966 NZ.CM00137 NC.018141 NC.013930 : 152 stability NC.014448 NC.016829 NC.017519 : 153 stability NC.010291 NC.010291 NC.010291 NC.010291 NC.01030 NC.017024 : 154 stability NC.007930 NC.017076 NC.017076 NC.015155 NC.015155 NC.015155 NC.015155 NC.015166 NC.015166 NC.011564 NC.011564 NC.011336 NC.011414 : 157 stability NC.008265 NC.007772 NC.007772 NC.007773 NC.011412 NC.015712	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. salivarius L. salivarius L. salivarius L. salivarius S. ruminantium measure : 0.476937 M. suis M. suis M. wenyonii measure : 0.734444 B. salanitronis B. salanitronis B. salanitronis C. canimorsus O. rhinotracheale measure : 0.931944 C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria Bacilli Bacilli <td< th=""><th>1.00 0.67 1.00 0.50 1.00 0.94 1.00</th></td<>	1.00 0.67 1.00 0.50 1.00 0.94 1.00
cluster : PL PL PL PL PL PL PL PL PL PL PL PL PL	138 stability NC.010721 NC.009433 NC.018830 NC.007617 NC.006139 NC.01314 NC.009739 NC.013284 NC.005240 NC.005240 NC.005240 NC.005240 NC.005241 139 stability NC.015171 NC.019149 NC.004986 NC.013776 140 stability NC.016751 NC.010003 141 stability NC.009488 NC.010793 142 stability NC.00843 NC.00843 NC.00843 NC.008509 NC.004843 NC.008511 NC.004343 NC.005824 NC.017552 NC.018021 143 stability NC.006912 144 stability NC.006912	measure : 0.799315 M. populi R. sphaeroides B. parapertussis N. multiformis D. psychrophila A. salmonicida C. sakazakii C. turicensis P. aeruginosa X. citri Y. pestis measure : 0.642879 S. hyicus S. aureus S. ruminantium Measure : 0.077862 M. piezophila P. mobilis measure : 0.473393 O. tsutsugamushi O. tsutsugamushi O. tsutsugamushi D. tsutsugamushi measure : 0.943750 L. biflexa L. borgpetersenii L. interrogans L. interrogans T. parva measure : 0.315348 N. farcinica 4. S. clavuligerus 9. S. clavuligerus Streptomyces sp. measure : 0.893333 B. cereus	αproteobacteria βproteobacteria βproteobacteria βproteobacteria γproteobacteria β γproteobacteria γproteobacteria β γproteobacteria φ γ <	0.92 1.00 0.87 1.00 1.00 0.95 0.91 0.93 0.93 0.93 0.93 1.00	PL CHR cluster PL PL PL PL CHR CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_009353 NC_008576 : 151 stability NC_014544 NC_009966 NZ_CM00137 NC_018141 NC_013930 : 152 stability NC_014448 NC_016829 NC_017519 : 153 stability NC_010291 NC_010291 NC_010291 NC_010291 NC_010330 NC_017024 : 154 stability NC_007930 NC_017076 NC_017076 NC_017076 NC_017078 : 155 stability NC_015155 NC_015155 NC_015156 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015166 NC_015167 NC_007772 NC_007773 NC_010937 NC_011412 NC_015712	M. gryphiswaldense M. gryphiswaldense M. marinus measure : 0.629402 L. longbeachae F. dumoffii 2 F. dumoffii L. pneumophila Thioalkalivibrio sp. measure : 0.638333 M. hyorhinis M. hyorhinis M. hyorhinis measure : 0.476937 E. faccium E. faccium E. faccium E. faccium E. faccium M. suivarius L. salivarius L. salivarius S. ruminantium S. ruminantium S. ruminantium M. suis M. suis D. rahintonis B. salanitronis B. salanitronis B. salanitronis B. salanitronis Cand. Azobacteroides C. canimorsus O. rhinotracheale measure : 0.931944 C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens C. perfringens	αproteobacteria Unidentified γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria %proteobacteria Mollicutes Clostridia Clostridia Clostridia Clostridia Clostridia Clostridia	$\begin{array}{c} 1.00\\ 0.67\\ \end{array}$

\mathbf{PL}	NC_008771	V. eiseniae	β proteobacteria	1.00
\mathbf{PL}	NC_014317	N. watsonii	γ proteobacteria	1.00
\mathbf{PL}	NC_014111	X. fastidiosa	γ proteobacteria	1.00
\mathbf{PL}	NC_014113	X. fastidiosa	γ proteobacteria	1.00
PL	NC_017561	X. fastidiosa	γ proteobacteria	1.00
cluster	: 159 stability	measure : 0.348431		
PL	NC_009806	K. radiotolerans	Actinobacteridae	0.82
\mathbf{PL}	NC_011355	M. liflandii	Actinobacteridae	0.88
\mathbf{PL}	NC_019018	M. marinum	Actinobacteridae	0.88
\mathbf{PL}	NC_005916	M. ulcerans	Actinobacteridae	0.82
cluster	: 160 stability	measure : 1.000000		
PL	NC_004429	E. coli	γ proteobacteria	1.00
PL	NC_010695	E. tasmaniensis	γ proteobacteria	1.00
PL	NC_002056	S. enterica	γ proteobacteria	1.00
PL	NC_009344	S. dysenteriae	γ proteobacteria	1.00
PL	NC_016834	S. sonnei	γ proteobacteria	1.00
aluatan	. 101 -t. Lilit.			
DI	NC 010402	Arthenhauter and		0.42
PL DI	NC 000478	C michiganonaia	Actinobacteridae	0.45
	NC 015664	Enankia symbiont	Actinobacteridae	0.12
PL DI	NC 017588	T therm embilies	Deinogagai	0.45
PL	NC 010878	Thermus SD	Deinococci	0.73
	NC 016624	Thermus sp.	Deinococci	1.00
1 1	110_010034	incinus sp.	Demococci	1.00
cluster	: 162 stability	measure : 0.109296		
PL	NC_008379	R. leguminosarum	α proteobacteria	1.00
\mathbf{PL}	NC_011371	R. leguminosarum	α proteobacteria	0.80
\mathbf{PL}	NC_012852	R. leguminosarum	α proteobacteria	1.00
		-	-	
cluster	: 163 stability	measure : 0.739940		
PL	NC_011774	B. cereus	Bacilli	1.00
\mathbf{PL}	NC_001988	$C.\ acetobuty licum$	Clostridia	1.00
PL	NC_015686	C. acetobutylicum	Clostridia	1.00
\mathbf{PL}	NC_017296	C. acetobutylicum	Clostridia	1.00
\mathbf{PL}	NC_014750	M. tractuosa	Cytophagia	1.00
cluster	: 164 stability	measure : 0.659632	<u> </u>	1.00
PL	NC_009466	C. kluyveri	Clostridia	1.00
PL	NC_011836	C. kluyveri E limi	Clostridia	0.57
PL	NC_019015	F. limi	Cytophagia	0.43
clustor	• 165 stability	measure \cdot 1 00000		
CHB	NC 016638	M haemocanis	Mollicutes	1.00
CHB	NC 014970	M haemofelis	Mollicutes	1.00
CHR	NC_017520	M. haemofelis	Mollicutes	1.00
CHR	NC_017520	M. haemofelis	Mollicutes	1.00
CHR cluster	NC_017520 : 166 stability	M. haemofelis measure : 0.738111	Mollicutes	1.00
CHR cluster PL	NC_017520 : 166 stability NC_006959	M. haemofelis measure : 0.738111 B. licheniformis	Mollicutes Bacilli	1.00
CHR cluster PL PL	NC_017520 : 166 stability NC_006959 NC_007706	M. haemofelis measure : 0.738111 B. licheniformis M. magneticum	Mollicutes Bacilli αproteobacteria	1.00 1.00 1.00
CHR cluster PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902	M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis	$\begin{array}{c} \text{Mollicutes} \\ \\ \text{Bacilli} \\ \alpha \text{proteobacteria} \\ \beta \text{proteobacteria} \end{array}$	1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903	M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis	Mollicutes Bacilli α proteobacteria β proteobacteria β proteobacteria	1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis	Mollicutes Bacilli α proteobacteria β proteobacteria β proteobacteria	1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL cluster	NC_007520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619	Mollicutes Bacilli α proteobacteria β proteobacteria β proteobacteria	1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL cluster PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551	M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli	1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL cluster PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli	1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL Cluster PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005200	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis L. lactis S. thermophilus G. divis	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL Cluster PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_008500	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis L. lactis S. thermophilus S. thermophilus	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL Cluster PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_008500 : 168 stability	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL cluster PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_008500 : 168 stability NC_010578	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL Cluster PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_008500 : 168 stability NC_015863	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline \end{array}$
CHR cluster PL PL PL PL PL PL PL Cluster PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_008500 : 168 stability NC_010578 NC_005863 NC_008741	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli βacilli βacilli βacilli βacilli	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_008500 : 168 stability NC_010578 NC_005863 NC_008863 NC_008741 NC_017311	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris	Mollicutes Bacilli α proteobacteria β proteobacteria Bacilli Bacilli Bacilli Bacilli α proteobacteria δ proteobacteria δ proteobacteria δ proteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL Cluster PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_0006959 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005500 : 168 stability NC_005863 NC_005863 NC_008741 NC_017311	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli βacilli βacilli βacilli βacilli βacilli βacilli βacilli βacilli βacilli βacilli βα	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL Cluster	NC_017520 : 166 stability NC_006959 NC_006959 NC_00706 NC_017902 NC_013551 NC_017495 NC_008500 : 168 stability NC_010578 NC_005863 NC_005863 NC_007311 : 169 stability	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris measure : 0.282287	Mollicutes Bacilli α proteobacteria β proteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli α proteobacteria δ proteobacteria δ proteobacteria δ proteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_00706 NC_010902 NC_010903 : 167 stability NC_013551 NC_013551 NC_005323 NC_005863 NC_005863 NC_005863 NC_007311 : 169 stability NC_004974	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. M. M	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Fusobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL Cluster PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_00706 NC_0170902 NC_010903 : 167 stability NC_013551 NC_013551 NC_013551 NC_005323 NC_008500 : 168 stability NC_005863 NC_008741 NC_017311 : 169 stability NC_004974 NC_002002	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. diatum F. nucleatum	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Fusobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1$
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_00700 NC_010902 NC_010903 : 167 stability NC_003551 NC_013551 NC_013551 NC_005323 NC_0058500 : 168 stability NC_005783 NC_005783 NC_005863 NC_0077311 : 169 stability NC_004974 NC_002002	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. diatum F. nucleatum	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Fusobacteria Fusobacteriia Fusobacteriia	$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ \hline 1.00\\ \hline 1.00\\ \hline 1.00\\ \hline 1.00\\ \hline 1.00\\ \hline \end{array}$
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 NC_017520 : 166 stability NC_006959 NC_007706 NC_0170902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005323 NC_005863 NC_008631 NC_008741 NC_002002 : 170 stability	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. vul	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Fusobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_006959 NC_007706 NC_017902 NC_013551 NC_013551 NC_005323 NC_008500 : 168 stability NC_005863 NC_005863 NC_005863 NC_004974 NC_002002 : 170 stability NC_008051 NC_0080551	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus S. thermophilus B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. coli	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Fusobacteria δproteobacteria δproteobacteria δproteobacteria εproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL Cluster PL PL PL Cluster PL PL	NC_017520 : 166 stability NC_017520 : 167 stability NC_00706 NC_01702 NC_010902 NC_013551 NC_017495 NC_005323 NC_005863 NC_005863 NC_005863 NC_0058741 NC_004974 NC_002002 : 170 stability NC_008051 NC_008052	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni	Mollicutes Bacilli αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli βproteobacteria δproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_00706 NC_00706 NC_010902 NC_010903 : 167 stability NC_013551 NC_013551 NC_013551 NC_005323 NC_005863 NC_005863 NC_005863 NC_007311 : 169 stability NC_004974 NC_008051 NC_008052 : 171 stability	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000	Mollicutes Bacilli αproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria	$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline \end{array}$
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_00706 NC_00706 NC_010902 NC_010903 : 167 stability NC_013551 NC_013551 NC_013551 NC_013551 NC_013532 NC_008500 : 168 stability NC_005863 NC_008741 NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. vulgaris measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli αproteobacteria δproteobacteria βproteobacteria δproteobacteria βproteobacteria βproteobacteria δproteobacteria δ	$\begin{array}{c} 1.00\\ \hline 1.00\\ \hline$
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL Cluster PL PL PL Cluster PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_007706 NC_0170902 NC_010903 : 167 stability NC_013551 NC_013551 NC_013551 NC_013551 NC_017495 NC_005823 NC_008500 : 168 stability NC_00863 NC_008631 NC_008051 NC_008052 : 171 stability NC_006822 NC_006145	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis M. magneticum C. taiwanensis measure : 0.747619 L. lactis L. lactis J. lactis S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. vulgaris M. Magneticum F. nucleatum F. nucleatum M. Magneticum M. Magnetic	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria βproteobacteria βροτοιο βροτοιο βροτοι <th>$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ \hline 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\$</th>	$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ \hline 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\$
CHR cluster PL PL PL PL PL PL PL PL PL PL Cluster PL PL PL Cluster PL PL PL Cluster PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_006959 NC_007706 NC_017902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005500 : 168 stability NC_005863 NC_005863 NC_007711 : 169 stability NC_004974 NC_008051 NC_008052 : 171 stability NC_006822 NC_006145	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides	Mollicutes Bacilli α proteobacteria β proteobacteria β proteobacteria Bacilli Bacilli Bacilli Bacilli α proteobacteria δ proteobacteria δ proteobacteria Fusobacteriia Fusobacteriia Encobacteria Encobacteria Bacilli Bacilli Bacilli	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_006959 NC_00706 NC_01750 NC_010902 NC_013551 NC_013551 NC_013551 NC_005323 NC_005863 NC_005863 NC_005863 NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability NC_006425 : 172 stability	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.00000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum M. measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000	Mollicutes Bacilli α proteobacteria β proteobacteria β proteobacteria Bacilli Bacilli Bacilli Bacilli Caproteobacteria δ proteobacteria δ proteobacteria Fusobacteriia Fusobacteriia Encobacteria ϵ proteobacteria ϵ proteobacteria ϵ proteobacteria	$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ \hline 1$
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL Cluster PL PL Cluster PL PL Cluster PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_00706 NC_01706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005863 NC_005863 NC_005863 NC_005863 NC_004974 NC_002002 : 169 stability NC_008051 NC_008052 : 171 stability NC_006145 : 172 stability NC_004941	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris M. vulgaris measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli Bacilli βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria βproteobacteria φroteobacteria βproteobacteria βproteobacteria φroteobacteria βproteobacteria βροτεοbacteria βροτεοbacteria βροτεοbacteria βροτεοbacteria βροτεοbacteria	$\begin{array}{c} 1.00\\ \hline 1.00\\ 1.00\\ \hline 1$
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 NC_017520 : 166 stability NC_006959 NC_00706 NC_010902 NC_010903 : 167 stability NC_013551 NC_013551 NC_013551 NC_013551 NC_013551 NC_013532 NC_005803 NC_005863 NC_005863 NC_005863 NC_004974 NC_002002 : 169 stability NC_004974 NC_008051 NC_008052 : 171 stability NC_006822 NC_006453 : 172 stability NC_004941 NC_004534	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli aproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria Fusobacteria Fusobacteria Bacilli Bacilli Actinobacteridae	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL Cluster PL PL PL Cluster PL PL Cluster PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_003532 NC_008500 : 168 stability NC_005863 NC_008741 NC_0077311 : 169 stability NC_008051 NC_008052 : 170 stability NC_0068052 : 171 stability NC_006845 NC_006453 : 172 stability NC_004941 NC_004534	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. vulgaris M. discum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli βproteobacteria δproteobacteria δproteobacteria Fusobacteriia Fusobacteria βproteobacteria βproteobacteria Actinobacteridae Actinobacteridae	$\begin{array}{c} 1.00\\$
CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_006959 NC_00706 NC_01796 NC_010902 NC_013551 NC_013551 NC_017495 NC_005323 NC_005863 NC_005863 NC_005863 NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability NC_006455 : 172 stability NC_004934 NC_004934	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis J. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum measure : 0.307077	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria βacilli Bacilli βproteobacteria δproteobacteria Fusobacteriia Fusobacteriia eproteobacteria eproteobacteria Bacilli Bacilli Actinobacteridae Actinobacteridae	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL Cluster PL PL cluster PL PL cluster PL PL cluster PL PL cluster PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_00706 NC_00706 NC_01702 NC_010902 NC_013551 NC_013551 NC_013551 NC_013551 NC_005323 NC-005863 NC_005863 NC_005863 NC_004974 NC_002002 : 169 stability NC.008051 NC_008052 : 171 stability NC.006822 NC-006415 : 172 stability NC.004941 NC-004534 : 173 stability	M. hacmofelis M. hacmofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum measure : 0.307077 N. gonorrhoeae	MollicutesBacilli α proteobacteria β proteobacteria β proteobacteriaBacilli Bacilli BacilliBacilli Bacilli α proteobacteria δ proteobacteria δ proteobacteriaFusobacteriia FusobacteriiaFusobacteriia eproteobacteria eproteobacteriaBacilli BacilliActinobacteridae Actinobacteridae β proteobacteria	1.00 1.00
CHR cluster PL PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL Cluster PL PL PL Cluster PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005803 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_008741 NC_017311 : 169 stability NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability NC_006453 : 172 stability NC_004941 NC_004534 : 173 stability NC_011034 NC_01034	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum M. gonorrhoeae A. pleuropneumoniae	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria Bacilli Bacilli Bacilli Bacilli βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL PL PL Cluster PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_00706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_003532 NC_008500 : 168 stability NC_005863 NC_008741 NC_005863 NC_008741 NC_002002 : 170 stability NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability NC_006822 NC_006145 : 172 stability NC_004941 NC_004534 : 173 stability NC_011034 NC_007800	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. magneticum M. and C. 200000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. augaris M.	MollicutesMollicutesBacilli α proteobacteria β proteobacteria β proteobacteriaBacilli Bacilli BacilliBacilli Bacilli α proteobacteria δ proteobacteria δ proteobacteriaFusobacteria FusobacteriaFusobacteria eproteobacteriaBacilli BacilliActinobacteridae Actinobacteridae β proteobacteria γ proteobacteria	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL Cluster PL PL Cluster PL PL cluster PL PL cluster PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_006959 NC_00706 NC_017902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005863 NC_005863 NC_005863 NC_004974 NC_004974 NC_008051 NC_008052 : 171 stability NC_006422 NC_004934 : 173 stability NC_004934 : NC_004934 : 173 stability NC_007800 NC_007800	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis J. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum M. measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum M. gonorrhoeae A. pleuropneumoniae B. trehalosi H. parasuis	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria βproteobacteria βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria βproteobacteria βproteobacteria </th <th>1.00 1.00</th>	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_017950 NC_010902 NC_013551 NC_013551 NC_017495 NC_005823 NC_005863 NC_005863 NC_005863 NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability NC_006822 NC_006455 : 172 stability NC_01034 : 173 stability NC_010889 NC_007800 NC_007800 NC_007800	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis J. lactis S. thermophilus S. thermophilus S. thermophilus D. vulgaris D. vulgaris D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. masure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum measure : 0.307077 N. gonorrhoeae A. pleuropneumoniae B. trehalosi H. parasuis measure : 0.57460	Mollicutes Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria βproteobacteria βproteobacteria βproteobacteria φproteobacteria βproteobacteria φproteobacteria βproteobacteria	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_004974 NC_004974 NC_004974 NC_004974 NC_004974 NC_008051 NC_008052 : 170 stability NC_006822 NC_006822 NC_006841 : 173 stability NC_011034 NC_01084 : 174 stability NC_006829 : 174 stability	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. vulgaris M. measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum measure : 0.307077 N. gonorrhoeae A. pleuropneumoniae B. trehalosi H. parasuis measure : 0.527460	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria Fusobacteriia Bacilli Bacilli Bacilli Bacilli Bacilli βproteobacteria φproteobacteria βproteobacteria β	1.00 1.00
CHR cluster PL PL PL PL cluster PL PL PL PL PL PL Cluster PL PL Cluster PL PL cluster PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_007706 NC_010902 NC_010903 : 167 stability NC_013551 NC_017495 NC_005323 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_005863 NC_008741 NC_017311 : 169 stability NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability NC_006425 NC_006453 : 172 stability NC_004941 NC_004941 NC_006829 : 174 stability NC_00889 NC_006829 : 174 stability NC_00811 NC_00811 NC_002111	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus M. measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. measure : 0.282287 F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum C. glutamicum M. gonorrhoeae A. pleuropneumoniae B. trehalosi H. parasuis measure : 0.527460 F. psychrophilum R. anatirestifer	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria βacilli Bacilli Bacilli βproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria δproteobacteria βproteobacteria β	1.00 1.00
CHR cluster PL PL PL PL PL PL PL PL PL PL Cluster PL PL Cluster PL PL Cluster PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC_017520 : 166 stability NC_006959 NC_00706 NC_00706 NC_010902 NC_010903 : 167 stability NC_013551 NC_013551 NC_013551 NC_013551 NC_005823 NC_005863 NC_005863 NC_008741 NC_002002 : 169 stability NC_004974 NC_002002 : 170 stability NC_008051 NC_008052 : 171 stability NC_006822 NC_006453 : 172 stability NC_004941 NC_004534 : 173 stability NC_007800 NC_006829 : 174 stability NC_004811 NC_002111	M. haemofelis M. haemofelis measure : 0.738111 B. licheniformis M. magneticum C. taiwanensis C. taiwanensis measure : 0.747619 L. lactis L. lactis S. thermophilus S. thermophilus measure : 1.000000 B. indica D. vulgaris D. vulgaris D. vulgaris D. vulgaris M. aulgaris measure : 0.282287 F. nucleatum F. nucleatum measure : 0.542348 C. coli C. jejuni measure : 1.000000 L. citreum L. mesenteroides measure : 1.000000 C. glutamicum C. glutamicum measure : 0.307077 N. gonorrhoeae A. pleuropneumoniae B. trehalosi H. parasuis measure : 0.527460 F. psychrophilum R. anatipestifer	Mollicutes Bacilli αproteobacteria βproteobacteria βproteobacteria βacilli Bacilli Fusobacteria φroteobacteria φroteobacteria βproteobacteria φroteobacteria βproteobacteria β β β β </th <th>1.00 1.00</th>	1.00 1.00

aluston	175 stability	monguno 1 0 706667		
PL	NC 014626	K wulaare	oproteobacteria	1.00
PL	NC_017385	K. vulgare	αproteobacteria	1.00
\mathbf{PL}	NC_018291	P. gallaeciensis	α proteobacteria	1.00
\mathbf{PL}	NC_018421	P. gallaeciensis	α proteobacteria	1.00
\mathbf{PL}	NC_015741	R. litoralis	α proteobacteria	1.00
cluster	• 176 stability	measure $\cdot 0.619222$		
PL	NC_009427	N. aromaticivorans	α proteobacteria	1.00
\mathbf{PL}	NC_015583	Novosphingobium sp.	α proteobacteria	1.00
cluster	177 stability	measure : 0.351816		1 00
PL	NC_011758	M. chloromethanicum	αproteobacteria	1.00
PL	NC 011892	M. extorquens M. nodulans	aproteobacteria	0.68
	1			1 0.00
cluster	: 178 stability	measure : 0.872222		
PL	NC_012219	C. botulinum	Clostridia	1.00
\mathbf{PL}	NC_012946	C. botulinum	Clostridia	1.00
PL	NC_017176	C. difficile	Clostridia	0.17
cluster	179 stability	measure $\cdot 0.293482$		
PL	NC 016586	A. lipoferum	αproteobacteria	1.00
PL	NC_013856	Azospirillum sp.	α proteobacteria	1.00
	1		•	
cluster	180 stability	measure : 1.000000		
PL	NC_011879	A. chlorophenolicus	Actinobacteridae	1.00
PL	NC_013325	S. aureus	Bacilli	1.00
гL	NG_013327	5. aureus	Bacilli	1.00
cluster	: 181 stability	measure : 0.081302		
PL	NC_009329	G. thermodenitrificans	Bacilli	0.38
\mathbf{PL}	NC_016646	Pseudovibrio sp.	α proteobacteria	0.00
\mathbf{PL}	NC_017624	S. enterica	γ proteobacteria	0.90
-l*	100			
DI	NC 010160	B tribocomum	oprotochastoria	0.72
PL	NC 009229	B. vietnamiensis	Broteobacteria	0.73
PL	NC_014723	B. rhizoxinica	β proteobacteria	0.71
\mathbf{PL}	NZ_CM00137	70 Desulfovibrio sp.	δ proteobacteria	1.00
cluster	: 183 stability	measure : 0.738111		1.00
PI.	NC 011961	S. thermophius T. roseum	Thermomicrobia	1.00
	1			1
clustor				
ciustei	: 184 stability	measure : 0.640119		
PL	184 stability NC_018582	measure : 0.640119 Gordonia sp.	Actinobacteridae	0.80
PL PL PL	NC_018583	measure : 0.640119 Gordonia sp. Gordonia sp.	Actinobacteridae Actinobacteridae	0.80
PL PL PL PL	: 184 stability NC_018582 NC_018583 NC_005307	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica	Actinobacteridae Actinobacteridae Actinobacteridae	0.80 1.00 1.00
PL PL PL cluster	 184 stability NC_018582 NC_018583 NC_005307 185 stability 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299	Actinobacteridae Actinobacteridae Actinobacteridae	$ \begin{array}{c} 0.80 \\ 1.00 \\ 1.00 \end{array} $
PL PL PL cluster CHR	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria	0.80 1.00 1.00
PL PL PL Cluster CHR CHR	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria	0.80 1.00 1.00 1.00
PL PL PL Cluster CHR CHR	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria	0.80 1.00 1.00 1.00
PL PL PL Cluster CHR CHR CHR	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798 : 186 stability	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria	0.80 1.00 1.00 1.00
Cluster PL PL PL CHR CHR CHR CHR PL PL PL	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798 : 186 stability NC_010378 NC_010954	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria	$ \begin{array}{c} 0.80 \\ 1.00 \\ 1.00 \\ 1.00 \\ 0.75 \\ 1.00 \\ \end{array} $
Cluster PL PL CHR CHR CHR CLuster PL PL PL	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798 : 186 stability NC_010378 NC_010378 NC_010954 NC_006856	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.80 \\ 1.00 \\ 1.00 \\ \end{array}$ $\begin{array}{c} 1.00 \\ 1.00 \\ 1.00 \\ \end{array}$
PL PL PL Cluster CHR CHR CHR CLuster PL PL PL	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798 : 186 stability NC_010378 NC_019054 NC_006856	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.80 1.00 1.00 1.00 0.75 1.00 1.00
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL Cluster	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798 : 186 stability NC_010378 NC_019054 NC_006856 : 187 stability	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.80 1.00 1.00 1.00 1.00 0.75 1.00 1.00
Cluster PL PL CHR CHR CHR CHR PL PL PL PL PL	: 184 stability NC_018582 NC_018583 NC_005307 : 185 stability NC_013009 NC_007798 : 186 stability NC_010378 NC_019054 NC_006856 : 187 stability NC_004951	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva	Actinobacteridae Actinobacteridae Δtinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.80 1.00 1.00 1.00 0.75 1.00 1.00 1.00
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_006856 187 stability NC_004951 NC_006088 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P. sputida	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	0.80 1.00 1.00 1.00 1.00 0.75 1.00 1.00 1.00
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL	 184 stability NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_006856 187 stability NC_004951 NC_006988 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp.	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria γproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL Cluster	 184 stability NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_019054 NC_006856 187 stability NC_004951 NC_006988 188 stability 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL CLUST CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_010378 NC_019054 NC_006856 187 stability NC_005009 NC_006988 188 stability NC_004058 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL CLUSTER CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_003007 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_010378 NC_019054 NC_006856 187 stability NC_004951 NC_005009 NC_006988 188 stability NC_004058 NC_004058 NC_004058 NC_004846 NC_004846 NC_004846 NC_004846 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria	0.80 1.00 1.00 1.00 0.75 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_010378 NC_010954 NC_006856 187 stability NC_004951 NC_005009 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_004846 NC_007206 NC_007206 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae Y. roctic	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL PL CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_010309 NC_007798 186 stability NC_010378 NC_010974 NC_010974 NC_006856 187 stability NC_004951 NC_006988 188 stability NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_007206 NC_015068 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae Y. pestis	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobact	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_010309 NC_007798 186 stability NC_010974 NC_0109754 NC_006856 187 stability NC_004051 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 Stability NC_004058 NC_004058 NC_004058 Stability NC_004058 Stability Stability 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria γproteobact	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$ $\begin{array}{c} 0.75\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$ $\begin{array}{c} 0.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC.018582 NC.018583 NC.005307 185 stability NC.013009 NC.007798 186 stability NC.010378 NC.019054 NC.004951 NC.004951 NC.004951 NC.004988 NC.004846 NC.007206 NC.015068 189 stability NC.015068 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria αproteobacteria γproteobacteria β	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_006856 187 stability NC_004951 NC_004951 NC_004951 NC_004958 188 stability NC_004058 NC_004058 NC_004058 188 stability NC_004058 189 stability NC_002191 NC_002191 NC_00909 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria γproteobacteria β	0.80 1.00
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_019054 NC_006856 187 stability NC_004951 NC_004951 NC_004951 NC_004958 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 I88 stability NC_004058 NC_004058 I89 stability NC_0015068 189 stability NC_002191 NC_014728 NC_014728 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P.seudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii L. delbrueckii L. delbrueckii	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria γproteobacteria βaproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018583 NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_004951 NC_004951 NC_004951 NC_004958 NC_004958 NC_004958 NC_004958 NC_004954 NC_004954 NC_004956 188 stability NC_004958 NC_004958	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii L. delbrueckii L. delbrueckii S. termitidis	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria γproteobacteria β β β <th>$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$</th>	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_010054 NC_000546 187 stability NC_004951 NC_004951 NC_004951 NC_005009 NC_004958 NC_004958 NC_004058 NC_004058 NC_004846 NC_004566 I89 stability NC_004568 I89 stability NC_015068 189 stability NC_013519 190 stability 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria γproteobacteria γproteobac	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_00798 186 stability NC_010378 NC_019054 NC_006856 187 stability NC_004951 NC_004951 NC_004951 NC_004951 NC_004953 NC_004954 NC_004956 NC_004956 NC_004951 NC_004956 NC_015068 189 stability NC_013519 190 stability NC_008610 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria aproteobacteria γproteobacteria γproteobact	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_010954 NC_006856 187 stability NC_004951 NC_004951 NC_004951 NC_004958 NC_004958 NC_004958 NC_004958 NC_004956 NC_015068 189 stability NC_002191 NC_0014728 NC_014728 NC_014729 NC_003610 NC_009465 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria aproteobacteria γproteobacteria γproteobact	0.80 1.00
Cluster PL PL PL PL CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_019054 NC_006856 187 stability NC_004951 NC_004951 NC_004058 NC_004058 NC_004846 NC_007206 NC_015068 189 stability NC_002191 NC_014728 NC_014728 NC_014751 S190 stability NC_008610 NC_009465 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P.seudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae aproteobacteria γproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL PL CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_004951 NC_0015068 189 stability NC_00465 191 stability NC_009465 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae aproteobacteria proteobacteria γproteobacteria γproteobact	0.80 1.00
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_006856 187 stability NC_004951 NC_004951 NC_004951 NC_004951 NC_004958 NC_004058 NC_004058 NC_004058 NC_004058 188 stability NC_0015068 189 stability NC_002191 NC_013519 190 stability NC_003465 191 stability NC_003465 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album	Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria γproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_006856 187 stability NC_004951 NC_004951 NC_004951 NC_004958 188 stability NC_004058 NC_007206 NC_015068 189 stability NC_002191 NC_002191 NC_002191 NC_013519 190 stability NC_008610 NC_013852 NZ_CM00147 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album N. halophilus	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae αproteobacteria γproteobacteria	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}\\ \hline \begin{array}{c} 0.75\\ 1.00\\ 1.00\\ \end{array}\\ \hline \begin{array}{c} 0.75\\ 1.00\\ 1.0$
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_004951 NC_004951 NC_004951 NC_004988 NC_004984 NC_004958 NC_004958 NC_0049846 NC_007206 NC_015068 189 stability NC_002191 NC_002191 NC_002191 NC_0014728 NC_013519 190 stability NC_008610 NC_003655 191 stability NC_013852 NZ_CM00147 NC_013858 NC_0014783 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album N. halophilus N. oceani	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria proteobacteria γproteobacteria γproteobac	0.80 1.00 0.95 0.97 0.53
Cluster PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_019054 NC_000856 187 stability NC_004951 NC_004951 NC_004951 NC_004951 NC_004951 NC_004953 NC_004954 NC_004954 NC_004954 NC_004955 188 stability NC_004956 189 stability NC_0015068 189 stability NC_002191 NC_010305 NC_003610 NC_009465 191 stability NC_013852 NZ_CM00147 NC_013854 NC_007483 NC_014316 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album N. halophilus N. oceani N. watsonii	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria aproteobacteria γproteobacteria γproteobac	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}\\ \hline 0.75\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.0\\ 1.$
Cluster PL PL PL PL PL CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018583 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_010378 NC_010054 NC_000954 NC_004951 NC_004951 NC_004951 NC_004951 NC_004958 NC_004958 NC_004058 NC_004058 NC_004058 NC_004058 NC_004058 NC_00491 NC_004058 NC_004846 NC_007466 NC_015068 189 stability NC_013651 NC_003610 NC_003852 NZ_CM00147 NC_013852 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P. putida P. seudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album N. halophilus N. oceani N. watsonii	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria aproteobacteria γproteobacteria γproteobacteria γproteobacteria </th <th>$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}\\ \hline 0.75\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00$</th>	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}\\ \hline 0.75\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00$
Cluster PL PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_00798 186 stability NC_019054 NC_004951 NC_004951 NC_004951 NC_004951 NC_004951 NC_004951 NC_004951 NC_004951 NC_004954 NC_004956 188 stability NC_004956 NC_004956 NC_015068 189 stability NC_002191 NC_013519 190 stability NC_003610 NC_003652 NZ_CM00147 NC_013958 NC_013958 NC_014316 192 stability NC_018971 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P. putida P.seudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album N. halophilus N. oceani N. watsonii measure : 0.362927	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria αproteobacteria γproteobacteria γproteobacte	0.80 1.00
Cluster PL PL PL PL PL CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_006856 187 stability NC_004951 NC_004951 NC_004951 NC_006988 188 stability NC_004968 NC_004846 NC_007206 NC_015068 189 stability NC_002191 NC_002191 NC_002191 NC_013519 190 stability NC_003552 NZ_CM00147 NC_013958 NC_007433 NC_014316 192 stability NC_00251 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P.seudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album N. halophilus N. oceani N. watsonii measure : 0.362927 R. jostii R. opacus	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria aproteobacteria γproteobacteria γproteobact	$\begin{array}{c} 0.80\\ 1.00\\ 1.00\\ 1.00\\ \end{array}\\ \hline 0.75\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00\\ \hline 1.00\\ 1.00\\ \hline 1.00$
Cluster PL PL PL PL CHR CHR CHR CHR PL PL PL PL PL PL PL PL PL PL PL PL PL	 184 stability NC_018582 NC_018583 NC_005307 185 stability NC_013009 NC_007798 186 stability NC_019054 NC_019054 NC_006856 187 stability NC_004951 NC_0042191 NC_013056 190 stability NC_003610 NC_003465 NZ_0014316 192 stability NC_008271 NC_013852 NZ_004351 	measure : 0.640119 Gordonia sp. Gordonia sp. G. westfalica measure : 0.510299 N. risticii N. sennetsu measure : 0.336294 E. coli E. coli S. enterica measure : 0.738826 P. fulva P. putida P. putida Pseudomonas sp. measure : 0.950000 H. influenzae H. influenzae H. influenzae H. influenzae H. influenzae Y. pestis measure : 0.797619 L. delbrueckii L. delbrueckii S. termitidis measure : 0.016676 C. Ruthia C. Vesicomyosocius measure : 0.467156 A. vinosum 76 M. album N. halophilus N. oceani N. watsonii measure : 0.362927 R. jostii R. opacus Rhodococcus sp.	Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteridae Actinobacteria aproteobacteria γproteobacteria γproteobact	0.80 1.00

cluster	: 193 stability measure : 1.000000		
PL	NC_006297 B. fragilis	Bacteroidia	1.00
\mathbf{PL}	NC_006873 B. fragilis	Bacteroidia	1.00
\mathbf{PL}	NC_015165 B. salanitronis	Bacteroidia	1.00
cluster	194 stability measure : 0.8666667		
BECE	NC 014388 B proteoclasticus	Clostridia	1.00
PL.	NC 014390 B proteoclasticus	Clostridia	1.00
1.0	1101014000 D. proteoclasticas	Clostifula	1.00
clustor	195 stability measure : 0 130478		
DI	NC 014924 D -thur	Cleataidia	1.00
	NC 014825 R -lbur	Clostridia	1.00
PL	NC_014825 R. alous	Clostridia	1.00
	100 (111)		
cluster	: 196 stability measure : 0.454697		1.00
PL	NC_019394 A. macleodii	γproteobacteria	1.00
PL	NC_015498 Glacieola sp.	γ proteobacteria	0.56
PL	NC_009705 Y. pseudotuberculosis	γ proteobacteria	0.59
cluster	: 197 stability measure : 0.332851		
\mathbf{PL}	NC_010509 M. radiotolerans	α proteobacteria	1.00
\mathbf{PL}	NC_010332 C. fungivorans	β proteobacteria	1.00
cluster	: 198 stability measure : 0.738826		
\mathbf{PL}	NC_013539 Planococcus sp.	Bacilli	1.00
\mathbf{PL}	NC_005076 S. sciuri	Bacilli	1.00
cluster	: 199 stability measure : 0.383412		
PL	NC_015423 A. denitrificans	β proteobacteria	1.00
\mathbf{PL}	NC_008765 Acidovorax sp.	β proteobacteria	1.00
cluster	200 stability measure : 0.643854		
PL	NC_004458 G. oxydans	α proteobacteria	1.00
PL	NC_010917 O. anthropi	α proteobacteria	1.00
\mathbf{PL}	NC_004965 S. meliloti	α proteobacteria	1.00
	1		
cluster	201 stability measure : 0.748593		
PL	NC_016623 A. lipoferum	αproteobacteria	1.00
\mathbf{PL}	NC_013857 Azospirillum sp.	αproteobacteria	1.00
cluster	202 stability measure : 0.021301		
PL.	NC 010795 A pleuroppeumoniae	oproteobacteria	nan
PL	NC 012661 H parasuis	oproteobacteria	1.00
12	110-012001 11. parasais	proteobacteria	1.00
cluster	203 stability measure : 0 525654		
DI	NC 018684 B thurin size size	Da ailli	1.00
PL	NC 010100 D. inuringiensis	Bacilli Da eilli	1.00
- 131			1 1 1 1 1
PL	NC_010180 B. weihenstephanensis	Bacim	1.00
	NC_010180 B. weihenstephanensis	Dacim	1100
PL	204 stability measure : 0.500680	Dacim	1.00
PL cluster PL	NC_010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC_012752 Cand. Hamiltonella	γproteobacteria	1.00
PL cluster PL PL PL	NC_010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC_012752 Cand. Hamiltonella NC_010899 V. cholerae	γ proteobacteria γ proteobacteria	1.00 1.00
PL cluster PL PL	NC_010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC_012752 Cand. Hamiltonella NC_010899 V. cholerae	γ proteobacteria γ proteobacteria	1.00 1.00
PL cluster PL PL cluster	NC_010180 B. weihenstephanensis 204 stability measure : 0.500680 NC_012752 Cand. Hamiltonella NC_010899 V. cholerae : 205 stability measure : 0.512387	γproteobacteria γproteobacteria	1.00 1.00
PL cluster PL PL cluster PL	NC_010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC_012752 Cand. Hamiltonella NC_010899 V. cholerae : 205 stability measure : 0.512387 NC_007412 A. variabilis	γproteobacteria γproteobacteria Homogonae	1.00 1.00 1.00
PL cluster PL PL cluster PL PL	NC_010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC_012752 Cand. Hamiltonella NC_010899 V. cholerae : 205 stability measure : 0.512387 NC_007412 A. variabilis NC_003240 Nostoc sp.	γproteobacteria γproteobacteria Homogonae Homogonae	1.00 1.00 1.00
PL cluster PL PL cluster PL PL	NC_010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC_012752 Cand. Hamiltonella NC_010899 V. cholerae : 205 stability measure : 0.512387 NC_007412 A. variabilis NC_003240 Nostoc sp.	γproteobacteria γproteobacteria Homogonae Homogonae	1.00 1.00 1.00 1.00
PL cluster PL PL cluster PL PL cluster	NC.010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720	γproteobacteria γproteobacteria Homogonae Homogonae	1.00 1.00 1.00 1.00
PL cluster PL PL cluster PL PL cluster PL	NC.010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC-014389 B. proteoclasticus	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia	1.00 1.00 1.00 1.00
PL cluster PL PL PL PL cluster PL PL PL	NC.010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00 0.57
PL cluster PL PL PL cluster PL PL PL PL	NC.010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00 0.57 1.00
PL cluster PL PL PL PL Cluster PL PL PL PL	NC.010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00
PL cluster PL PL PL PL cluster PL PL PL cluster	NC.010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.015688 C. acetobutylicum NC.012657 C. botulinum : 207 stability measure : 0.697000	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia	1.00 1.00 1.00 1.00 1.00 0.57 1.00
PL cluster PL PL PL cluster PL PL PL cluster PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC-005026 NC.005026 B. fragilis	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia	1.00 1.00 1.00 1.00 1.00 0.57 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 B. fragilis NC.002132 Flavobacteria sp.	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Flavobacteriia	1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95
PL cluster PL PL PL cluster PL PL PL Cluster PL PL PL PL	NC.010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.014155 T. intermedia	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteriia βproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ \hline 1.00\\ 0.57\\ 1.00\\ \hline 1.00\\ 0.95\\ 1.00\\ \hline \end{array}$
PL cluster PL PL cluster PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.005026 B. fragilis NC.002132 Flavobacteria sp. NC.014155 T. intermedia	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Flavobacteriia βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL cluster	NC.010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.014155 T. intermedia : 208 stability measure : 0.195886	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteriia βproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009350 A. salmonicida	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteria βproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\ 1.00\\ 0.95\\ 1.00\\ 0.83\\ \end{array}$
PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL	NC.010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.014155 T. intermedia : 208 stability measure : 0.195886 NC.009350 A. salmonicida	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteria βproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\ 0.95\\ 1.00\\ 0.83\\ 1.00\\ \end{array}$
PL cluster PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.005122 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009352 NC.009352 A. salmonicida NC.009352 V. fischeri	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 B. fragilis NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009350 NC.009350 A. salmonicida NC.006842 V. fischeri	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteriia βproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95 1.00 1.00 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL cluster	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.002132 NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009350 NC.009350 A. salmonicida NC.009352 A. salmonicida NC.006842 V. fischeri 209 stability measure : 0.567660	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Bacteroidia Flavobacteriia βproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95 1.00 0.83 1.00
PL cluster PL PL PL cluster PL PL PL cluster PL PL PL PL PL cluster PL PL	NC.010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.0014155 T. intermedia : 208 stability measure : 0.195886 NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009352 A. salmonicida NC.006842 V. fischeri : 209 stability measure : 0.567660 NC-004952 F. novicida	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Bacteroidia Flavobacteriia βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\ 0.95\\ 1.00\\ 0.83\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
PL cluster PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.005026 B. fragilis NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009352 A. salmonicida NC.009352 A. salmonicida NC.009352 A. salmonicida NC.009452 F. novicida NC.004952 F. novicida	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteriia γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	1.00 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL cluster PL PL PL cluster PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009350 NC.00352 A. salmonicida NC.006842 V. fischeri 209 stability measure : 0.567660 NC.004952 NC.013092 F. philomiragia NC.013092 F. philomiragia NC.013092 F. philomiragia	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pavobacteria βproteobacteria γproteobacteria	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010809 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.012657 C. botulinum Stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.0014155 T. intermedia : 208 stability measure : 0.195886 NC.009350 A. salmonicida NC.009352 A. salmonicida NC.006842 V. fischeri : 209 stability measure : 0.567660 NC.004952 F. novicida NC.013092 F. philomiragia NC.002109 F. tularensis	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\ 0.95\\ 1.00\\ 0.83\\ 1.00\\$
PL cluster PL PL PL cluster PL PL PL cluster PL PL PL PL PL Cluster PL PL PL PL cluster PL PL PL Cluster	NC.010180 B. weinenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.0014155 T. intermedia : 208 stability measure : 0.195886 NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009352 F. novicida NC.004952 F. novicida NC.004952 F. novicida NC.002109 F. tularensis : 210 stability measure : 0.804648	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\ 0.95\\ 1.00\\ 0.83\\ 1.00\\$
PL cluster PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.000302 NC.005026 B. fragilis NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009352 NC.009352 A. salmonicida NC.009352 A. salmonicida NC.009352 F. novicida NC.004952 F. novicida NC.003092 F. philomiragia NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 R. albus	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Clostridia	1.00 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009350 NC.00352 A. salmonicida NC.006842 V. fischeri 209 stability measure : 0.567660 NC.004952 NC.0014025 F. novicida NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 R. albus NC.018680 A. macleodii	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Proteobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\ 1.00\\ 1.00\\ 0.95\\ 1.00\\$
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis : 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010809 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. protocolasticus NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009350 A. salmonicida NC.00842 V. fischeri : 209 stability measure : 0.567660 NC.004952 F. novicida NC.002109 F. tularensis : 210 stability measure : 0.804648 NC.014827 <td< th=""><th>γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria</th><th>$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$</th></td<>	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$
PL cluster PL PL cluster PL PL PL cluster PL PL PL Cluster PL PL PL PL Cluster PL PL PL Cluster PL PL PL Cluster	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.002132 NC.001312 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009350 NC.009350 A. salmonicida NC.009352 A. salmonicida NC.004952 F. novicida NC.002109 F. tularensis 209 stability measure : 0.567660 NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 R. albus NC.018680 A. macleodii : 211 stability measure : 0.642879	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Bacteroidia Flavobacteria γproteobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$
PL cluster PL PL cluster PL PL PL cluster PL PL Cluster PL PL Cluster PL PL PL cluster PL PL PL Cluster PL PL PL Cluster	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.002132 Flavobacteria sp. NC.0014155 T. intermedia 207 stability measure : 0.195886 NC.0020352 NC.009350 A. salmonicida NC.009352 A. salmonicida NC.009352 F. novicida NC.002109 F. tularensis 210 stability measure : 0.567660 NC.014827 R. albus NC.014827 R. albus NC.018680 A. macleodii 211 stability measure : 0.642879 NC_003528 L. reuteri	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria β β β γproteobacteria β β γproteobacteria β	1.00 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 B. fragilis NC.002132 Flavobacteria sp. NC.014155 T. intermedia 208 stability measure : 0.195886 NC.009350 NC.00352 A. salmonicida NC.004532 F. novicida NC.004952 F. novicida NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 R. albus NC.014827 R. albus NC.014827 R. albus NC.018680 A. macleodii 211 stability measure : 0.642879 NC.003528 L. reuteri	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria βproteobacteria β β β	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\ 1.00\\ 0.95\\ 1.00\\$
PL cluster PL PL cluster PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.014389 B. proteoclasticus NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.0020132 Flavobacteria sp. NC.014155 T. intermedia : 208 stability measure : 0.195886 NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009350 A. salmonicida NC.009352 F. novicida NC.004952 F. novicida NC.004952 F. novicida NC.002109 F. tularensis : 210 stability measure : 0.804648 NC.014827 R. albus NC.013680 A. macleodii : 211 stability measure : 0.642879	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria β β β γproteobacteria γproteobacteria β β γproteobacteria β β γproteobacteria β	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL cluster PL PL PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.005026 B. fragilis NC.002132 Flavobacteria sp. NC.014155 T. intermedia 207 stability measure : 0.195886 NC.002132 Flavobacteria sp. NC.004155 T. intermedia 208 stability measure : 0.195886 NC.009350 A. salmonicida NC.009352 A. salmonicida NC.004952 F. novicida NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 NC.014827 R. albus NC.003528 L. reuteri NC.004532	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria β β β γproteobacteria β β γproteobacteria β	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ \end{array}$
PL cluster PL PL cluster PL PL PL PL cluster PL PL PL Cluster PL PL PL PL PL cluster PL PL PL PL PL Cluster PL PL PL PL PL PL	NC.010180 B. weinenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.010809 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.002132 Flavobacteria sp. NC.014155 T. intermedia 207 stability measure : 0.195886 NC.009350 NC.009350 A. salmonicida NC.009352 A. salmonicida NC.004952 F. novicida NC.004952 F. novicida NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 NC.018680 A. macleodii 211 stability measure : 0.642879 NC.003528 L. reuteri NC.003528 L. reuteri NC.003528 L. reuteri	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Sproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria β β γproteobacteria β γproteobacteria β β γproteobacteria γproteobacteria β	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.57 1.00 1.00 0.95 1.00 0.83 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010809 V. cholerae : 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. : 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.012657 C. botulinum : 207 stability measure : 0.697000 NC.002132 Flavobacteria sp. NC.002132 Flavobacteria sp. NC.002132 Flavobacteria sp. NC.002132 Flavobacteria : 207 stability measure : 0.195886 NC.009350 NC.002132 Flavobacteria NC.009350 A. salmonicida NC.009352 A. salmonicida NC.004952 F. novicida NC.014952 F. novicida NC.014927 R. albus NC.014827 R. albus NC.014827 R. albus NC.014827 R. albus NC.014827 R. albus NC.003528 L. reuteri <th>γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria %proteobacteria %proteobacteria</th> <th>1.00 1.00</th>	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria %proteobacteria	1.00 1.00
PL cluster PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.014389 B. proteoclasticus NC.012657 C. botulinum 207 stability measure : 0.697000 NC.002026 NC.002132 Flavobacteria sp. NC.014155 T. intermedia 207 stability measure : 0.195886 NC.009350 NC.009350 A. salmonicida NC.009352 A. salmonicida NC.009352 F. novicida NC.004952 F. novicida NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 R. albus NC.014827 R. albus NC.003528 L. reuteri NC.003528 L. reuteri NC.004532 L. reuteri NC.005511	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria proteobacteria proteobacteria Sproteobacteria Sphingobacteria Sphingobacteria Sphingobacteria	1.00 1.00
PL cluster PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster PL PL PL cluster	NC.010180B. weihenstephanensis: 204 stability measure : 0.500680NC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.012899V. cholerae: 205 stability measure : 0.512387NC.007412A. variabilisNC.003240Nostoc sp.: 206 stability measure : 0.193720NC.014389B. proteoclasticusNC.015688C. acetobutylicumNC.012657C. botulinum: 207 stability measure : 0.697000NC.005026B. fragilisNC.002132Flavobacteria sp.NC.014155T. intermedia: 208 stability measure : 0.195886NC.009352A. salmonicidaNC.009352F. novicidaNC.009352F. novicidaNC.002109F. tularensis: 209 stability measure : 0.567660NC.004952F. novicidaNC-014827R. albusNC-014827R. albusNC-018680A. macleodii: 211 stability measure : 0.642879NC-003528L. reuteri: 212 stability measure : 0.215810NC-015511H. hydrossisNC-015512H. hydrossisNC-015512H. hydrossis	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria βproteobacteria γproteobacteria βproteobacteria γproteobacteria βproteobacteria γproteobacteria βproteobacteria βριηροbacteria βριηροbacteria βριηροbacteria	1.00 1.00
PL cluster PL PL Cluster PL PL PL PL PL PL Cluster PL PL Cluster PL PL PL PL Cluster PL PL PL PL PL Cluster PL PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL	NC.010180B. weihenstephanensis: 204 stability measure : 0.500680NC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.010899V. cholerae: 205 stability measure : 0.512387NC.007412A. variabilisNC.003240Nostoc sp.: 206 stability measure : 0.193720NC.014389B. protocolasticusNC.012657C. botulinum: 207 stability measure : 0.697000NC.002026B. fragilisNC.002132Flavobacteria sp.NC.014155T. intermedia: 208 stability measure : 0.195886NC.009350A. salmonicidaNC.009352A. salmonicidaNC.009352F. novicidaNC.004952F. novicidaNC.014952F. novicidaNC.012109F. tularensis: 209 stability measure : 0.567660NC.004952F. novicidaNC.001897R. albusNC-018680A. macleodii: 211 stability measure : 0.642879NC.004532L. reuteri: 212 stability measure : 0.215810NC-015511H. hydrossis: 213 stability measure : 0.738111NC 015218L. oridaptibut	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria βproteobacteria Sphoteobacteria Sphingobacteria Sphingobacteria Sphingobacteria Sphingobacteria	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 1.00\\ 0.57\\ 1.00\\$
PL cluster PL PL cluster PL PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL	NC.010180B. weihenstephanensis: 204 stability measure : 0.500680NC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.010899V. cholerae: 205 stability measure : 0.512387NC.007412A. variabilisNC.003240Nostoc sp.: 206 stability measure : 0.193720NC.014389B. proteoclasticusNC.012657C. botulinum: 207 stability measure : 0.697000NC.002132Flavobacteria sp.NC.014155T. intermedia: 208 stability measure : 0.195886NC.009350A. salmonicidaNC.009350A. salmonicidaNC.009350A. salmonicidaNC.009352F. novicidaNC.004952F. novicidaNC.004952F. novicidaNC.002109F. tularensis: 210 stability measure : 0.567660NC.014827R. albusNC.01580A. macleodii: 211 stability measure : 0.642879NC.003528L. reuteri: 212 stability measure : 0.215810NC.015511H. hydrossis: 213 stability measure : 0.738111NC.01528L. acidophilusNC.015218L. acidophilusNC.01522Stability measure : 0.738111	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Sphingobacteria Sphingobacteria Sphingobacteria Bacilli Bacilli Bacilli	1.00 1.00
PL cluster PL PL cluster PL PL cluster PL PL cluster PL PL Cluster PL PL cluster PL PL cluster PL PL cluster PL PL cluster PL PL cluster PL PL cluster PL PL Cluster PL PL Cluster PL PL Cluster PL PL PL PL PL PL PL PL PL PL PL PL PL	NC.010180 B. weihenstephanensis 204 stability measure : 0.500680 NC.012752 Cand. Hamiltonella NC.012752 Cand. Hamiltonella NC.010899 V. cholerae 205 stability measure : 0.512387 NC.007412 A. variabilis NC.003240 Nostoc sp. 206 stability measure : 0.193720 NC.014389 B. proteoclasticus NC.015688 C. acetobutylicum NC.012657 C. botulinum 207 stability measure : 0.697000 NC.005026 NC.005122 Flavobacteria sp. NC.014155 T. intermedia 207 stability measure : 0.195886 NC.00352 A. salmonicida NC.009350 A. salmonicida NC.009352 F. novicida NC.004952 F. novicida NC.002109 F. tularensis 210 stability measure : 0.804648 NC.014827 R. albus NC.003528 L. reuteri NC.003528 L. reuteri NC.004532 L. reuteri NC.005511 H. hydrossis NC.015512 H. hydrossis <	βacini γproteobacteria γproteobacteria Homogonae Clostridia Clostridia Clostridia Bacteroidia Flavobacteriia βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$
PL cluster PL PL cluster PL PL PL PL cluster PL PL cluster PL PL Cluster PL PL Cluster PL PL Cluster PL PL Cluster PL PL Cluster PL PL Cluster PL PL PL Cluster PL PL PL PL PL PL PL PL PL PL	NC.010180B. weinenstephanensis: 204 stability measure : 0.500680NC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.010899V. cholerae: 205 stability measure : 0.512387NC.007412A. variabilisNC.003240Nostoc sp.: 206 stability measure : 0.193720NC.014389B. proteoclasticusNC.015688C. acetobutylicumNC.012657C. botulinum: 207 stability measure : 0.697000NC.005026B. fragilisNC.002132Flavobacteria sp.NC.014155T. intermedia: 208 stability measure : 0.195886NC.009350A. salmonicidaNC.009352A. salmonicidaNC.009352A. salmonicidaNC.009352A. salmonicidaNC.002109F. tularensis: 209 stability measure : 0.567660NC.004952F. novicidaNC.013092F. philomiragiaNC.012109F. tularensis: 210 stability measure : 0.804648NC.014827R. albusNC.01580A. macleodii: 211 stability measure : 0.642879NC.003528L. reuteri: 212 stability measure : 0.215810NC.015511H. hydrossis: 213 stability measure : 0.738111NC.015218L. acidophilusNC.015218L. acidophilusNC.015224L. amylovorus	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Proteobacteria βproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria βproteobacteria βacilli Bacilli Bacilli Bacilli Bacilli	$\begin{array}{c} 1.00\\ 1.00\\ 1.00\\ 1.00\\ \hline \end{array}$
PL cluster PL PL cluster PL PL PL PL PL PL PL PL PL PL	NC.010180B. weihenstephanensis: 204 stability measure : 0.500680NC.012752Cand. HamiltonellaNC.0012752Cand. HamiltonellaNC.012752Cand. HamiltonellaNC.010899V. cholerae: 205 stability measure : 0.512387NC.007412A. variabilisNC.003240Nostoc sp.: 206 stability measure : 0.193720NC.014389B. proteoclasticusNC.012657C. botulinum: 207 stability measure : 0.697000NC.002132Flavobacteria sp.NC.014155T. intermedia: 208 stability measure : 0.195886NC.009350A. salmonicidaNC.009350A. salmonicidaNC.009350A. salmonicidaNC.009352A. salmonicidaNC.009352F. novicidaNC.004952F. novicidaNC.002109F. tularensis: 210 stability measure : 0.567660NC.014827R. albusNC.018680A. macleodii: 211 stability measure : 0.642879NC.003528L. reuteriNC.004532L. reuteriNC.015511H. hydrossis: 213 stability measure : 0.215810NC.015218L. acidophilusNC.015322L. amylovorus: 214 stability measure : 0.738111NC.0157NC.01572NC0157NC0157NC0157	γproteobacteria γproteobacteria Homogonae Homogonae Clostridia Clostridia Clostridia Clostridia Pacteroidia Flavobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria γproteobacteria Sphoteobacteria Sphingobacteria Bacilli Bacilli Bacilli Bacilli Bacilli Bacilli	1.00 1.00

\mathbf{PL}	NC_001772	Clostridium sp.	Clostridia	1.00
cluster	: 215 stability	measure : 0.248163		
PL	NC_015695	R. slithyformis	Cytophagia	1.00
1.5	10-013/31	D. IIngulate	Oytophagia	0.55
cluster	: 216 stability	measure : 1.000000		1.00
PL PL	NC_017070	S. ruminantium	Negativicutes	1.00
12	110-011011	5. Tantilantiant	regativicates	1.00
cluster	: 217 stability	measure : 0.382924	A 1	1.00
PL PL	NC_002699	Frankia sp. Nocardionsis sp.	Actinobacteridae	1.00
	1			
cluster	: 218 stability	measure : 0.347563		1.00
PL PL	NC_008770 NC_017284	C. jejuni C. iejuni	εproteobacteria εproteobacteria	1.00
	1		-p	
cluster	: 219 stability	measure : 0.733050	Classication in	1.00
PL PL	NC_005231 NC_009931	Synechocystis sp. A. marina	Chroobacteria Homogonae	1.00
cluster	: 220 stability	measure : 1.000000	Antinabantaridaa	1.00
PL	NZ_CM00101	.6 S. clavuligerus	Actinobacteridae	1.00
	1	0		1
cluster	221 stability	measure : 0.672639	Questo al antonia	1.00
CHR	NC_017293	Cand. Tremblaya	βproteobacteria	1.00
	1	U	, .	I
cluster	222 stability	measure : 0.232208	Actinobactoridao	0.20
CHR	NC_008148	R. xylanophilus	Rubrobacridae	0.20
\mathbf{PL}	NC_018606	Cardinium symbiont	Bacteroidetes	0.55
PL CHB	NC_019012	F. aestuarina A thermophila	Cytophagia Anaerolineae	1.00
CHR	NC_017079	C. aerophila	Caldilineae	0.67
CHR	NC_014314	D. lykanthroporepellens	Dehalococcoidetes	0.57
PL PL	NC_010182	B. weihenstephanensis L. sphaericus	Bacilli Bacilli	0.53
PL	NC_010078	W. cibaria	Bacilli	0.36
PL	NC_018066	D. acidiphilus	Clostridia	0.27
PL	NC_006177	S. thermophilum F. nucleatum	Fusobacteriia	0.73
CHR	NC_012489	$G. \ aurantiaca$	Gemmatimonadetes	0.64
CHR	NC_011296	T. yellowstonii	Nitrospira	0.63
PL	NC_016596	A. brasilense	αproteobacteria	0.64
PL	NC_016618	A. brasilense	α proteobacteria	0.68
PL PL	NC_013855	Azospirillum sp. M. radiotolerans	αproteobacteria	0.00
PL	NC_017959	T. mobilis	α proteobacteria	0.53
CHR	NC_005363	B. bacteriovorus	δ proteobacteria	0.86
CHR	NC_014365 NC 015388	D. baarsii D. acetoxidans	oproteobacteria δproteobacteria	0.00
CHR	NC_011768	D. alkenivorans	δ proteobacteria	0.75
CHR	NC_018025	D. tiedjei	δ proteobacteria	0.43
CHR	NC_008554	S. cellulosum S. fumaroxidans	δproteobacteria δproteobacteria	$0.00 \\ 0.12$
PL	NC_016850	A. fischeri	γ proteobacteria	0.00
PL	NC_017318	E. tarda M. adhaarana	γ proteobacteria	0.27
PL	NC_014838	Pantoea sp.	γproteobacteria	0.73
PL	NC_014840	Pantoea sp.	γ proteobacteria	1.00
CHB CHB	NC_014841	Pantoea sp. T violascens	γproteobacteria γproteobacteria	0.64
PL	NC_010614	V. tapetis	γ proteobacteria	0.64
CHR	NC_015681	T. indicus	Thermodesulfobacteria	0.17
CHR	NC_015682	Thermodesulfobacterium sp.	Thermodesulfobacteria	0.36
CHR	NC_010794	M. infernorum	Verrumicrobia	1.00
\mathbf{CHR}	NC_010655	A. muciniphila	Verrucomicrobiae	1.00

<u>Résumé</u> :

Le génome bactérien est classiquement pensé comme constitué de "chromosomes", éléments génomiques essentiels pour l'organisme, stables et à évolution lente, et de "plasmides", éléments génomiques accessoires, mobiles et à évolution rapide. La distinction entre plasmides et chromosomes a récemment été mise en défaut avec la découverte dans certaines lignées bactériennes d'éléments génomiques intermédiaires, possédant à la fois des caractéristiques de chromosomes et de plasmides. Désignés par le terme de "chromosomes secondaires", "mégaplasmides" ou "chromid", ces éléments sont dispersés parmi les lignées bactériennes et sont couramment décrits comme des plasmides adaptés et modifiés. Cependant, leur véritable nature et les mécanismes permettant leur intégration dans le génome stable reste à caractériser. En utilisant les protéines liées aux Systèmes de Transmission de l'Information Génétique (STIG) comme variables descriptives des éléments génomiques bactériens (ou réplicons), une étude globale de génomique comparative a été conduite sur l'ensemble des génomes bactériens disponibles. À travers l'analyse de l'information contenue dans ce jeu de données par différentes approches analytiques, il apparaît que les STIG constituent des marqueurs pertinents de l'état d'intégration des réplicons dans le génome stable, ainsi que de leur origine évolutive, et que les Réplicons Extra-Chromosomiques Essentiels (RECE) témoignent de la diversité des mécanismes génétiques et des processus évolutifs permettant l'intégration de réplicons dans le génome stable, attestant ainsi de la continuité du matériel génomique.

$\underline{Mots-clé}$:

génome bactérien, réplicon, systèmes de transmission de l'information génétique (STIG), fouille de données, apprentissage automatique, classification, analyses multivariées, discrimination fonctionnelle, synténie, chromosome, plasmide, réplicon extra-chromosomique essentiel (RECE), néochromosome.

Summary :

The genome of bacteria is classically separated into essential, stable and slow evolving replicons (chromosomes) and accessory, mobile and rapidly evolving replicons (plasmids). This paradigm is being questioned since the discovery of extra-chromosomal essential replicons (ECERs), be they called "megaplasmids", "secondary chromosomes" or "chromids", which possess both chromosomal and plasmidic features. These ECERs are found in diverse lineages across the bacterial phylogeny and are generally believed to be modified plasmids. However, their true nature and the mechanisms permitting their integration within the sable genome are yet to be formally determined. The relationships between replicons, with reference to their genetic information inheritance systems (GIIS), were explored under the assumption that the inheritance of ECERs is integrated to the cell cycle and highly constrained in contrast to that of standard plasmids. A global comparative genomics analysis including all available of complete bacterial genome sequences, was performed using GIIS functional homologues as parameters and applying several analytical procedures. GIIS proved appropriate in characterizing the level of integration within the stable genome, as well as the origins, of the replicons. The study of ECERs thus provides clues to the genetic mechanisms and evolutionary processes involved in the replicon stabilization into the essential genome and the continuity of the genomic material.

Keywords :

bacterial genome, replicon, genetic information inheritance systems (GIIS), data mining, machine learning, classification, multivariate analyses, functional discrimination, synteny, chromosome, plasmid, extra-chromosomal essential replicon (ECER), neo-chromosome.